

(12)特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2003年11月6日 (06.11.2003)

PCT

(10) 国際公開番号
WO 03/091872 A1

(51) 国際特許分類⁷:

G06F 7/24

(21) 国際出願番号:

PCT/JP03/05319

(22) 国際出願日: 2003年4月25日 (25.04.2003)

(25) 国際出願の言語:

日本語

(26) 国際公開の言語:

日本語

(30) 優先権データ:

特願2002-125838 2002年4月26日 (26.04.2002) JP

(71) 出願人(米国を除く全ての指定国について): 学校法人日本大学 (NIHON UNIVERSITY SCHOOL JURIDICAL PERSON) [JP/JP]; 〒102-8275 東京都千代田区九段南四丁目8番24号 Tokyo (JP).

(72) 発明者; および

(75) 発明者/出願人(米国についてのみ): 山本 登 (YAMAMOTO,Noboru) [JP/JP]; 〒963-8846 福島県郡山市久留米五丁目106-2 Fukushima (JP).

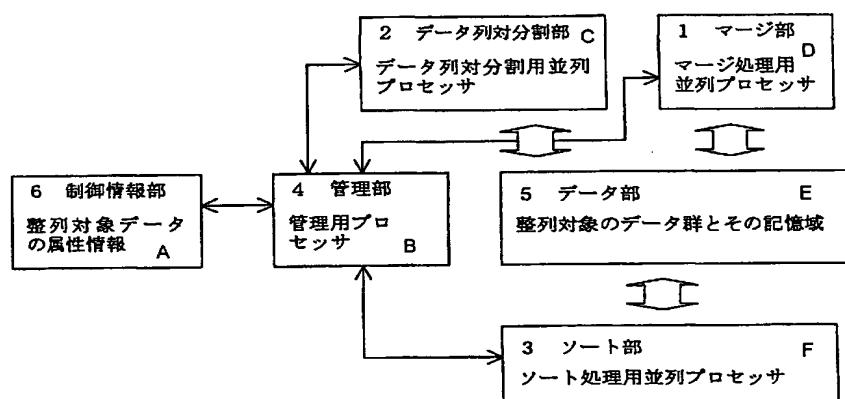
(74) 代理人: 大木 健一 (OHKI,Kenichi); 〒113-0034 東京都文京区湯島二丁目15番1-205号 Tokyo (JP).

(81) 指定国(国内): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

[統葉有]

(54) Title: PARALLEL MERGE/SORT PROCESSING DEVICE, METHOD, AND PROGRAM

(54) 発明の名称: 並列マージソート処理装置及び方法並びにプログラム



- A...6 CONTROL INFORMATION UNIT ATTRIBUTE INFORMATION ON DATA TO BE SORTED
- B...4 MANAGEMENT UNIT PROCESSOR FOR MANAGEMENT
- C...2 DATA STRING PAIR DIVISION UNIT PARALLEL PROCESSOR FOR DIVIDING DATA STRING PAIR
- D...1 MERGE UNIT PARALLEL PROCESSOR FOR MERGE PROCESSING
- E...5 DATA UNIT DATA GROUPS TO BE SORTED AND THEIR STORAGE AREAS
- F...3 SORT UNIT PARALLEL PROCESSOR FOR SORT PROCESSING

(57) Abstract: In the basic form of merge processing, one of a sort processing, one sorted data string is output as a whole on receiving a pair of two sorted partial data strings. Conventionally, a high parallelism of this processing has been considered to be difficult. A method for dividing a sorted partial data string pair into a plurality of segment pairs, if invented, would allow a high-level of merge processing to be performed even in a homogeneous-configuration parallel computer system such as a tightly-coupled multiprocessor sharing a main storage. The basic of merge processing is to receive a pair of two sorted partial data string and to output one sorted data string. A method for sub-dividing an input data string pair into arbitrary data string pairs from the first part of the input data string pair while considering the magnitude of the key value has been invented. This method, if executed, enables a data

string pair to be merged at a parallelism k if the data string is divided into any number of data string pairs. When merging a data string pair in descending order (merging it from the first part thereof to the last part and outputting it from the first part to the last part in the output area) and in ascending order (merging it from the last part to the first part and outputting it from the last part to the first part in the output area), the merging can be at parallelism 2k.

[統葉有]



(84) 指定国(広域): ARIPO 特許 (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), ヨーラシア特許 (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ特許 (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI 特許 (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:
— 國際調査報告書

2 文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

(57) 要約:

整列処理のひとつであるマージ処理の基本形では、2つの整列済部分データ列の対を入力とし、全体として整列済の1系列のデータ列を出力する。従来この操作の高並列化は困難とされていた。整列済部分データ列対を複数の区画対に分割する方法があれば、主記憶を共有する密結合マルチプロセッサのような均質な構成の並列計算機システムでも高度な並列マージ処理が可能となる。マージ処理の基本は2つの整列済部分データ列の対を入力し一つの整列済データ列を出力する処理である。この入力データ列対を両データ列の先頭からキー値の大小を考慮して任意のデータ列対に細分割する方法を発明した。この方法を実現すると、任意数のデータ列対、例えば k 組のデータ列対に分割すると並列度 k でのマージ操作が可能となるほか、正順併合(データ列対の頭部から尾部に向かって併合し出力域の頭部から尾部に向かって出力する)と逆順併合(データ列対の尾部から頭部に向かって併合し出力域の尾部から頭部に向かって出力する)を適用すれば並列度 $2k$ のマージ操作も可能となる。

明細書

並列マージソート処理装置及び方法並びにプログラム

5 技術分野

この発明の核心となる技術は、2つの整列済データ列を入力とするマージソート処理を任意数の並列度でできるように、2つの整列済データ列対（つい）をキー値の大小を考慮して複数の整列済部分データ列対に細分割する技術である。

それを用いた実施例として並列2分木マージソート法を取り上げる。まず未整列データ列を利用可能なプロセッサの数に対応する数に分割し並列にソート操作を行い、並列数だけの整列済部分データ列を生成する。次に、2組の整列済部分データ列を対にしてマージ操作を行うのだが、処理速度を高めるために、入力データ列対を必要な数だけのデータ列対に細分割し、それぞれ並列にマージ操作を行う。この出力は細分割せずに1つのマージ操作をした時と同じ結果になるよう各マージ操作の実行時パラメータで出力域を指定する。本発明は、この操作を整列された単一のデータ列を得る迄繰り返し行うことを特徴とする並列マージソート処理装置及び方法並びにプログラムに関するものである。

背景技術

コンピュータを用いた基本的な処理の1つにデータの整列がある。整列（ソート）処理及び併合（マージ）処理は、入力されたデータ列を着目するフィールドの情報に関して昇順（小から大へ）又は降順（大から小へ）に整列して1系列のデータ列を得る操作である。併合操作は昇順（または降順）に整列された2系列以上のデータ（整列対象のデータまたはレコードの総称）列を受けて、昇順または降順に整列された1系列のデータ列を出力する。

入力系列の数は、2、3、4・・・と種々の場合が考えられるが、特に2の場合は1回の比較で25 1単位のデータの整列位置が確定するため効率が良い。

併合（マージ）処理に係る所要時間のオーダはデータ数(n)に比例するため、 n^2 または $n \log_2 n$ に比例するソート操作に比べ有利である。しかしデータ数の増大と共に所要時間も増すので、より高速なアルゴリズム(算法)が求められている。また性能の高いアルゴリズムの開発により、マージソートの適用範囲がさらに広まる可能性もある。だが2つのキー値の比較という単純な操作

であるため、時間短縮の要因が潜在する可能性は低い。本発明は従来技術ではなし得なかった 2 つの整列済部分データ列対を入力とするマージ処理の並列化を可能とするものである。

並列プロセッサを用いることにより多数のデータに対して同時並行的にマージ処理を行い、処理時間を短縮することが提案されている。しかしプロセッサ間の接続ネットワークに特殊なトポ 5 ロジーを要するものや特殊な機能・構造のプロセッサを要するものが多く、主記憶を共有する密結合マルチプロセッサのような汎用的な計算機システムでは実現が困難である。本発明は構成や接続に特殊性を持たない汎用的な並列プロセッサで高い並列性をもつマージソート法を実現するものだが、そのような方法を用いないと非常に効率が低い処理方式の例を以下に示す。

汎用的な並列プロセッサシステムに適用可能な並列マージソート方法の一例として並列 2 分木 10 マージソート算法をあげる。図 1 3 はその概略フローチャートであり、図 1 4 はプロセッサ数が 8 の場合の処理説明図である。

図 1 3 を参照すると、処理対象の未整列データ列（入力データ）をプロセッサの数に対応する p 組のデータ列（ただし、 $p = 2^q$ 、q は整数）に分割する（S 1 0 0）。p 個のプロセッサを用 15 い、分割された p 組の未整列データ列を例えばクイックソート法によりそれぞれ独立かつ並行して整列する（S 1 0 1）ことができる。整列された p 組の部分データ列に対して q 段のマージ処理を施す（S 1 0 2）ことにより、最終的に全体として整列された 1 組のデータ列を得ることができる（S 1 0 3）。

図 1 4 を参照してプロセッサの数が 8 の場合の処理について説明する。図中、丸はプロセッサを示し、四角はデータ列あるいはデータ列（D）が記憶される領域を示す。丸の中の記号はプロ 20 セッサが行う処理の内容を示し、S は整列（ソート）処理、M は併合（マージ）処理、V は必要に応じてなされる他の記憶領域への転送処理を意味する。

入力領域に置かれた未整列のデータ列は 8 つの部分データ列に分割される。8 つのプロセッサ P 1 ~ P 8 によりこれら部分について同時並行的にソート処理を行う。これらの処理結果は領域 D 1 1 ~ D 1 8 に置かれる。

次に、第 1 段のマージ処理を行う。プロセッサ P 1 によりデータ列 D 1 1 と D 1 2 が併合され 25 て領域 D 2 1 に置かれ、プロセッサ P 3 によりデータ列 D 1 3 と D 1 4 が併合されて領域 D 2 2 に置かれ、プロセッサ P 5 によりデータ列 D 1 5 と D 1 6 が併合されて領域 D 2 3 に置かれ、プロセッサ P 7 によりデータ列 D 1 7 と D 1 8 が併合されて領域 D 2 4 に置かれる。なお、第 1 段においてプロセッサ P 2, P 4, P 6, P 8 は使用されない（プロセッサの割り当ては便宜上の 30 ものである。以下同じ）。

次に、第2段のマージソートを行う。プロセッサP1によりデータ列D21とD22が併合されて領域D31に置かれ、プロセッサP5によりデータ列D23とD24が併合されて領域D32に置かれる。なお、第2段においてプロセッサP2, P3, P4, P6, P7, P8は使用されない。

5 次に、第3段のマージソートを行う。プロセッサP1によりデータ列D31とD32が併合されて領域D4に置かれる。これでマージソートが完了する。なお、第3段においてプロセッサP2, P3, P4, P5, P6, P7, P8は使用されない。得られた結果D4は必要なら8つのプロセッサにより最終出力領域に転送される。

10 上述した従来の手法では、マージ処理段が進行につれ使われなくなるプロセッサが増加するためプロセッサの処理能力が無駄になっていた。上記の例では第1段のマージ処理ではプロセッサの50%しか使用されず、第2段のマージ処理のプロセッサ使用率は25%、第3段のマージ処理ではプロセッサ使用率は12.5%となり、処理段が進むにつれて使用されないプロセッサの数が増える。これはマージ処理段がすすむにつれて整列済部分データ列の数が半減していくためである。

15 また、1台のプロセッサでマージするデータ数は後段になるに従って倍増するため、処理時間が増大する。高速処理を優先するシステムでは、プロセッサの使用率低下の問題以上に、1台のプロセッサが扱うデータ量の増大による処理時間の増大が問題となる。

20 この発明は、係る課題を解決するためになされたもので、並列プロセッサを用いたマージソート処理においてプロセッサの使用効率を高め、処理時間を短縮可能な並列マージソート処理装置及び方法並びにプログラムを提供することを目的とする。

発明の開示

(段落0014) 本発明ではマージ処理の入力となる2つの整列済部分データ列の対を任意数のデータ列対に細分割して並列にマージ処理をすることにより、マージ処理の所要時間を短縮するものである。複数のプロセッサから共通にデータ群が参照されることから、主記憶を共有する複数のプロセッサを接続した並列プロセッサにおいて特に有効であるが、データ群が各プロセッサから参照できる環境にある如何なる並列プロセッサにも適用可能である。

並列2分木マージソート処理では、未整列の入力データ列を複数組に分割し、それぞれについてソート処理を行って整列済部分データ列を作る。以後は存在する整列済部分データ列を2組ず30 つ対にしてマージ処理を行い1組の整列済部分データ列を出力する処理段を、出力データ列の数

が1つにまとまる迄繰り返すものである。各マージ処理段では入力となるデータ列対が複数ある限りマージ処理は並列に行われる。

1 単位のジョブあるいはタスクであるマージ処理の並列化は従来技術では不可能なため、マージ処理段の進行と共に並列度は半減し処理時間は倍増する。

5 本発明によれば、マージ処理の入力となる2組の整列済部分データ列（図1-4に示したソート処理やマージ処理によって得られた部分的な整列済データ列）の対を複数の整列済部分データ列の対に細分割することにより細分割した数だけのプロセッサを使用することができ、整列処理に要する時間を短縮できる。

また、並列度を高めるため以下に記載する方法を適用することもできる。すなわち、入力となる2組の整列済データ列の対に対し、2つのプロセッサを用いて、キー値の大きい側（始端）からマージ操作を始めて結果を出力域の始端から降順に出力する方法（正順併合または正順マージ）と、キー値の小さい側（終端）からマージ操作を始めて出力域の終端から昇順に出力する方法（逆順併合または逆順マージ）の両方のマージ操作を同時平行的に実行するのである。

この方式自体に起因する制約はないため、複数のタスクが並行に動作できる環境下では容易に実現できる。

以下の説明においては、説明の便宜上『一つのキーによる降順の整列』に限定する。整列に関連する用語には原則として『整列』と『併合』を用いるが、同義語的に『ソート』や『マージ』を用いることがある。

本発明に係る並列マージソート処理装置は、従来方式のマージ操作における2系列の入力データ列の対を複数組の2系列の入力データ列対に分割する機構と、それらの組に対して並列にマージ操作を行う機構を基本要素とする。

両機構に対しどのようにプロセッサ群を割り付けるかは任意である。すなわち、データ列対の細分割を複数のプロセッサにさせた後、マージ処理では改めてプロセッサを割り当て直すこともできるし、データ列の細分割操作とその後のマージ操作を同一プロセッサに担当させることもできる。

それぞれのマージ操作の出力はあたかも1つのマージ操作を行った場合と同じ結果でなければならない。このため、並行に動作する各マージ操作では、細分割した入力データ列対全体のキー値の大小に対応した領域に結果を出力する必要がある。

部分データ列の対を割り当てられた前記プロセッサ群は、それぞれマージ処理を行うと共に当該マージ処理の結果を予め指示された領域に出力する。このためには、例えばマージ処理の入力

データ列対が格納されている領域の先頭メモリアドレスの対と処理すべきデータ数と出力域の先頭アドレスが管理プロセッサから指定されればよい。

前記部分データ列の対それぞれに対して2つのプロセッサが割り当てられ、第1のプロセッサは前記部分データ列の一方の端（キー値の大きい側）から正順にマージ処理を行い、第2のプロセッサは同じ部分データ列の他方の端（キー値の小さい側）から逆順にマージ処理を同時に行い、前記第1のプロセッサは前記部分データ列の一方の端から順にマージ処理の結果を書き込み、前記第2のプロセッサは前記領域の部分の他方の端から順にマージ処理の結果を書き込む。

(段落0020) 段落0014に記述したように1つのデータ列対を複数のデータ列対に細分するが、細分割は指定されたデータ列対の始点から始めて指定数xのデータ数が存在するデータ列と残りのデータ列とに2分する操作をxを変化させて反復する操作に分解できる。

従って細分割の基本はデータ列の2分操作であるが、その結果得られる2つの整列済部分データ列対が満たすべき条件は以下のように定式化される。

n個のデータからなる整列済データ列Dを(D, n)で表わし、2つのデータ列を対にしたものを{(D₁, n), (D₂, n)}と表記する。

整列済のデータ列の対:{(D₁, n), (D₂, n)}が、2つの部分データ列{(D₁₁, n₁₁), (D₂₁, n₂₁)}と、部分データ列{(D₁₂, n₁₂), (D₂₂, n₂₂)}に分割され、それらがマージ処理に使われて全体としてキー値の大小に矛盾のない出力をするためには下式が成立しなければならない。

部分データ列D₁₁とD₂₁の末尾のデータのキー値の小さい方

≥ 部分データ列D₁₂とD₂₂の頭部のデータのキー値の大きい方

$n_{11} + n_{21} = 2$ 分した上位データ列対に含まれるべきデータ数

この発明に係る並列マージソート処理方法は、複数のプロセッサを含む並列プロセッサを用いて並列マージソート処理を行うための方法であって、

入力として、未整列の1つのデータ列、または2つの整列済部分データ列、あるいは3つ以上の整列済部分データ列のいずれかとその属性情報（種別・データの所在情報や個数など）を受け入れ、

整列対象が未整列データ列の場合はそのデータ列を分割して複数の未整列部分データ列を得る

第1ステップと、

前記複数のソート用部分データ列にそれぞれプロセッサを割り当てる第2ステップと、割り当てられたプロセッサにより前記複数の部分データ列についてそれぞれ独立にソート処理を行い降順に整列された部分データ列を出力する第3ステップと、

5 第3ステップまたは第7ステップで得たか初期データとして入力された、任意の2つの整列済部分データ列でマージソート用の入力データ列の対をつくり、所定の分割条件の下でそれぞれ要求された数の細分された部分データ列の対に分割する第4ステップと、

細分した部分データ列対に対してマージ処理をするために必要なジョブ情報を編集する第5ステップと、

各マージ処理に対してプロセッサを割り当てる第6ステップと、

10 割り当てられたプロセッサによりマージ処理を行い整列された部分データ列を出力する第7ステップと、

出力データ列の数が1つになる迄前記第4ステップ乃至前記第7ステップを繰り返すステップとを備えるものである。

但し、以上は1つの未整列データ列が与えられる場合に適用されるもので、2つ以上の整列済データ列が与えられる場合は第1ステップから第3ステップ迄は不要である。

また第4ステップ乃至前記第7ステップの繰り返しはマージ処理された整列済部分データ列が1つのデータ列になった段階で終わる。特に2つの整列済データ列が与えられる場合は第4ステップから第7ステップは1回だけ行われ繰り返しは必要としない。

好ましくは、

20 前記第6ステップにおいて、前記部分データ列の対に対して2つのプロセッサが割り当てられ、前記第7ステップにおいて、第1のプロセッサは前記部分データ列の一方の端（キー値の大きい側）から正順マージ処理を行い、第2のプロセッサは同じ部分データ列の他方の端（キー値の小さい側）から逆順マージ処理を同時に行う。

（2分した下位データ列対の頭部の位置決定）

25 段落0020に記したデータ列対の2分方法を以下に記述する。

◆前記データ列D₁とD₂内の個々のデータを指定する指標変数iとjを設け、それらへの初期値として前記xを設定（10000個ずつのデータ列対を1000個ずつのデータ列対10区画に分割する場合、個数としてのx=1000、分割位置の初期指標には1000, ··· 9000を設定）するステップと、

◆データ列D₁の指標変数iが指すデータのキー値とデータ列D₂の指標変数jが指すデータのキー値を比較する比較ステップと、

◆最初の比較でD₁の指標変数iが指すデータのキー値とD₂の指標変数jが指すデータのキー値が等しくない場合、キー値の大きい方の指標変数に1を加え、小さい方の指標変数から1を引いた後、前記比較ステップに分岐するステップと、

◆2回目以降の比較でD₁の指標変数iが指すデータのキー値とD₂の指標変数jが指すデータのキー値との大小関係に変化がない場合、大きい方の指標変数に1を加え、小さい方の指標変数から1を引いた後、前記比較ステップに分岐するステップと、

◆最初の比較でD₁の指標変数iが指すデータのキー値とD₂の指標変数jが指すデータのキー値とが等しいとき、指標変数iが指すデータと指標jが指すデータをそれぞれ分割境界とするステップと、

◆D₁の指標変数iが指すデータのキー値とD₂の指標変数jが指すデータのキー値との間の大小関係が前回の大小関係から反転（2回目以降の比較では等しい場合も反転とみなす）したとき、前回の比較操作におけるD₁のキー値とD₂のキー値のうちの大きい方と、今回の比較操作におけるD₁のキー値とD₂のキー値の内の大きい方とを比較して、小さい方のキー値をもつデータを分割境界とすると共に、当該データの当初の比較相手を他方の境界とするステップ

とを実行する。なおこの処理は2分した後半のデータ列対{(D₁₂, n₁₂), (D₂₂, n₂₂)}の頭部の決定、換言すれば正順併合の開始点の決定に使われる。

（2分した上位データ列対の尾部の位置決定）

20 段落0020に記したように2分する方法を以下に記述する。

◆前記データ列D₁とD₂の個々のデータを指定する指標変数iとjとに境界の指標値の初期値として前記x-1を設定（10000個ずつのデータ列対を1000個ずつのデータ列対10区画に分割する場合、個数としてのxには1000、分割位置の初期指標には999, 1999, .. 8999を設定する）するステップと、

25 ◆データ列D₁の指標変数iが指すデータのキー値とデータ列D₂の指標変数jが指すデータのキー値を比較する比較ステップと、

◆最初の比較でD₁の指標変数iが指すデータのキー値とD₂の指標変数jが指すデータのキー値が等しくない場合、キー値の大きい方の指標変数に1を加え、小さい方の指標変数から1を引いた後、前記比較ステップに分岐するステップと、

◆2回目以降の比較でD₁の指標変数iが指すデータのキー値とD₂の指標変数jが指すデータのキー値との大小関係に変化がない場合、大きい方の指標変数に1を加え、小さい方の指標変数から1を引いた後、前記比較ステップに分岐するステップと、

◆最初の比較でD₁の指標変数iが指すデータのキー値とD₂の指標変数jが指すデータのキー値とが等しいとき、指標変数iが指すデータと指標jが指すデータをそれぞれ分割境界とするステップと、

◆D₁の指標変数iが指すデータのキー値と領域D₂の指標変数jが指すデータのキー値との間の大小関係が前回の大小関係から逆転（2回目以降の比較では等しい場合も反転とみなす）したとき、前回の比較操作におけるD₁のキー値とD₂のキー値のうちの小さい方と、今回の比較操作におけるD₁のキー値とD₂のキー値の内の小さい方とを比較して、大きい方のキー値をもつデータを分割境界とすると共に、当該データの当初の比較相手を他方の境界とするステップ

とを実行する。なおこの処理は2分した前半のデータ列対{(D₁₁, n₁₁), (D₂₁, n₂₁)}の尾部の決定、換言すれば逆順併合の開始点の決定に用いる。

この発明は、マージソートの入力となる2組の整列済データ列の対をキー値の大小分布を考慮して複数のデータ列対に細分割することにより、任意の並列度のマージソート処理を可能とする並列マージソート処理方法であって、

整列されたデータからなる第1のデータ列と第2のデータ列を用意するステップと、

要求される並列度に応じて前記第1のデータ列及び前記第2のデータ列の対を両データ列の頭部からのキー値の大小を考慮して複数の整列済部分データ列対に細分割するステップと、分割された整列済部分データ列対を入力として並列にマージ処理を行うステップとを備えるものである。

この発明は、複数のプロセッサを含む並列プロセッサを用いて並列マージソート処理を行うためのプログラムであって、

入力として、未整列の1つのデータ列、または2つの整列済部分データ列、あるいは3つ以上の整列済部分データ列のいずれかとその属性情報（種別・データの所在情報や個数など）を受け入れ、

整列対象が未整列データ列の場合はそのデータ列を分割して複数の未整列部分データ列を得る第1ステップと、

前記複数のソート用部分データ列にそれぞれプロセッサを割り当てる第2ステップと、

割り当てられたプロセッサにより前記複数の部分データ列についてそれぞれ独立にソート処理を行い降順に整列された部分データ列を出力する第3ステップと、

第3ステップまたは第7ステップで得たか初期データとして入力された、任意の2つの整列済部分データ列でマージソート用の入力データ列の対をつくり、所定の分割条件の下でそれぞれ要求された数の細分された部分データ列の対に分割する第4ステップと、

5 細分した部分データ列対に対してマージ処理をするために必要なジョブ情報を編集する第5ステップと、

各マージ処理に対してプロセッサを割り当てる第6ステップと、

割り当てられたプロセッサによりマージ処理を行い降順に整列された部分データ列を

各マージ処理に対してプロセッサを割り当てる第6ステップと、

割り当てられたプロセッサによりマージ処理を行い降順に整列された部分データ列を出力する

10 第7ステップと、

マージ処理されたデータ列を前記部分データ列として前記第4ステップ乃至前記第7ステップを出力データ列の数が1つになる迄繰り返すステップとを備えるものである。

但し第1ステップから第3ステップ迄は入力データ列が未整列の場合にのみ必要なものである。

また第4ステップ乃至前記第7ステップの繰り返しはマージ処理された整列済部分データ列が1
15 つのデータ列になった段階で終わる。特に2つの整列済データ列が与えられる場合は第4ステップから第7ステップは1回だけ行われ繰り返しは必要としない。

この発明に係るプログラムは、例えば、記録媒体に記録される。

媒体には、例えば、E P R O Mデバイス、フラッシュメモリデバイス、フレキシブルディスク、
20 ハードディスク、磁気テープ、光磁気ディスク、CD (C D - R O M、V i d e o - C Dを含む)、
DVD (D V D - V i d e o、D V D - R O M、D V D - R A Mを含む)、R O Mカートリッジ、
バッテリバックアップ付きのR A Mメモリカートリッジ、フラッシュメモリカートリッジ、不揮
発性R A Mカートリッジ等を含む。

また、電話回線等の有線通信媒体、マイクロ波回線等の無線通信媒体等の通信媒体を含む。イ
ンターネットもここでいう通信媒体に含まれる。

25 媒体とは、何等かの物理的手段により情報（主にデジタルデータ、プログラム）が記録されて
いるものであって、コンピュータ、専用プロセッサ等の処理装置に所定の機能を行わせることが
できるものである。要するに、何等かの手段でもってコンピュータにプログラムをダウンロード
し、所定の機能を実行させるものであればよい。

30 図面の簡単な説明

図 1 は、発明の実施の形態に係る並列マージソート処理装置の概要を示す図である。

図 2 は、発明の実施の形態 1 に係る処理フローチャートである。

図 3 は、発明の実施の形態に係る処理全体の説明図である。

図 4 は、発明の実施の形態に係る処理の説明図である。

5 図 5 は、発明の実施の形態に係る処理の説明図である。

図 6 は、発明の実施の形態に係る処理の説明図である。

図 7 は、発明の実施の形態に係る処理の一部の説明図である。

図 8 は、発明の実施の形態に係る領域分割のフローチャートである（正順併合）。

図 9 は、発明の実施の形態に係る領域分割のフローチャートである（逆順併合）。

10 図 10 は、発明の実施の形態に係る領域分割の説明図である。

図 11 は、発明の実施の形態に係る領域分割の例を示す図である。

図 12 は、発明の実施の形態 2 に係る処理フローチャートである。

図 13 は、従来の並列マージソート処理のフローチャートである。

図 14 は、従来の並列マージソート処理の説明図である。

15

発明を実施するための最良の形態

発明の実施の形態 1.

発明の実施の形態に係る装置／方法について図面を参照して説明する。

図 1 は本装置の概略構成を示すものである。1 は複数のプロセッサを含み、複数のタスクを同時に並行的に実行可能なマージ処理用並列プロセッサである。2 は2組の整列済データ列の対を所定数の部分データ列対に分割して管理用プロセッサ 4 に返す領域分割用並列プロセッサである。3 は指定された領域に存在する未整列のデータ列をソートするソート処理用並列プロセッサである。また 4 は全体の制御を担当する管理プロセッサで、ジョブ開始に際して外部から受ける入出力データの場所やデータ数などの情報 6 を受けるほか、前記並列プロセッサ群に行わせるタスクに必要な情報の編集やタスクの起動、タスク終了報告の受け付けなどの処理を担当する。5 は整列対象のデータを含むデータ記憶域である。

1 と 3 は実装によっては同じ並列プロセッサを用いることもできるし、2 のデータ列対分割用並列プロセッサについても同様である。

従来技術では一つのマージ処理用としてしか扱われない2つの整列済データ列の対を複数の部分データ列の対に分割することにより並列度を向上させる。後述の図 3 の例では、マージの入力

となるデータ列D 2 1とD 2 2の対を2分割することにより2組の入力データ列対を得て、それを正順マージと逆順マージを行うことによりプロセッサP 1～P 4に同時に処理をさせることができ、データ列D 3 1、D 3 2の対を4分割することにより4組の入力データ列対を得て、それぞれを正順マージと逆順マージを行うことによりプロセッサP 1～P 8に同時に処理をさせ

5 ることができる。詳しくは後述する。

図2の概略フローチャートで示す並列2分木マージソート法では、整列の対象であるデータ列を $p = 2^a$ の関係をもつp組の未整列のデータ列に分割し、それぞれを独立に、例えばクイックソート法により整列させる。

S 1～S 3は入力データ列を複数のデータ列に分割し p 個のプロセッサに同時並行的にソート処理を行わせるための手順である。その後、以下のように q 段のマージソートを繰り返し、最終的に全体として整列された1組のデータ列を得る。

S 1：整列処理の対象である未整列のデータ列を p 組の未整列データ列に分割する。ここで、 $p = 2^a$ で、動員可能なプロセッサの数に相当し、 q はマージの段数を表す数である。短い処理時間を探る場合は p を大きくするが、 $\log_2 p$ の関係で q も増加する。

15 また説明の便宜上用いるマージ段を表す変数 a を0に初期化する。

S 2： p 個のプロセッサを分割された p 区画に割り当てる。

S 3：それぞれ独立にソートする。

S 4：マージ操作の入力となる整列済部分データ列の対を分割する。

(1) $a = a + 1$

20 (2) 2^{a-a+1} 組の整列済部分データ列を2つずつ組み合わせた 2^{a-a} 組の入力データ列対を作る。

各入力データ列対のそれぞれを 2^{a-1} 組の細分化されたデータ列対に分割する。データ列対の分割の仕方については後述する。

S 5： $(2^{a-a}) \times (2^{a-1}) = 2^{a-1}$ 組の細分されたデータ列対のそれぞれに正順併合と逆順併合用のプロセッサを割り当てる。

S 6：各組についてそれぞれ2つのプロセッサを用いて正順併合と逆順併合を同時に 2^{a-a} 組の整列された部分データ列を得る。

S 7：次に、 $a \neq q$ なら a に1を加え、S 4～S 7の処理を繰り返す。

S 8：以上の手順により、最終的に全体として整列された1組のデータ列を得ることができる。

図2の方法は、1段目のマージでは分割しないが、2段目において2分割し、3段目において4分割し、4段目において8分割し、・・・、q段目において領域を 2^{q-1} 分割し、それぞれの領域に正順併合と逆順併合を適用する。図2の整列済データ列分割法を適用した並列2分木マージソート法では常にp並列の併合操作を行うことができる。

5 図3を参照してプロセッサの数が8の場合における並列2分木マージソート法の処理について説明する。図中、丸はプロセッサを示し、四角はデータ列あるいはデータ列(D)が記憶される領域を示す。丸の中の記号はプロセッサが行う処理の内容を示し、Sは整列(ソート)処理、Mは併合(マージ)処理、Vは転送処理を意味する。

10 入力領域に置かれたデータは8組の未整列データ列に分割する。8つのプロセッサP1～P8によりこれら部分データ列について同時並行的にソート処理を行う。これらの処理結果は領域D11～D18に置かれる。

15 次に、第1段目のマージソートを行う。プロセッサP1、P2によりデータ列D11とD12が併合されてD21に置かれ、プロセッサP3、P4によりデータ列D13とD14が併合されてD22に置かれ、プロセッサP5、P6によりデータ列D15とD16が併合されてD23に置かれ、プロセッサP7、P8によりデータ列D17とD18が併合されてD24に置かれる。

なお第1段目では正順併合と逆順併合が同時に行われるので、分割を行わなくても8台のプロセッサが使用される(プロセッサの割り当ては便宜上のものである。以下同じ)。

20 次に、第2段目のマージソートを行う。全てのプロセッサを使用するためにデータ列対を2分割する。すなわち、部分データ列D21とD22の対に対しきー値の大小を考慮して2分割し、D21-1とD22-1の対とD21-2とD22-2の対を作る。データ列D23とD24の対についても同様である。

25 そして、プロセッサP1、P2によりデータ列D21-1とD22-1が併合されてD31の一方(キー値の大きい側)に置かれ、プロセッサP3、P4によりデータ列D21-2とD22-2が併合されてD31の他方(キー値の小さい側)に置かれる。同様に、プロセッサP5～P8によりデータ列D23とD24が併合されてD32に置かれる。

30 次に、第3段目のマージソートを行う。第2段階目と同様に分割が行われる。すなわち、データ列D31とD32を対にし、キー値の大小を考慮して4分割し、D31-1とD32-1、D31-2とD32-2、D32-3とD32-3、D31-4とD32-4の各部分データ列対を作る。プロセッサP1、P2によりデータ列D31-1とD32-1が併合されてD4内の所定の位置に置かれ、プロセッサP3、P4によりデータ列D31-2とD32-2が併合されて

領域D 4 内の所定の位置に置かれ、プロセッサ P 5、P 6 によりデータ列D 3 1－3 と D 3 2－3 が併合されて領域D 4 内の所定の位置に置かれ、そして、プロセッサ P 7、P 8 によりデータ列D 3 1－4 と D 3 2－4 が併合されて領域D 4 内の所定の位置に置かれる。これでマージソートが完了する。得られた結果D 4 は必要なら8台のプロセッサにより出力領域に転送される。

5 図3の処理の各段階において8つのプロセッサの全てが使用される。

図4～図6を参照して図3の処理を模式的に説明する。

d10～d13、d20～d23 の8つの部分データ列（各ブロックのデータ数=380個）の合計3040組のデータをソートすることを考える。

10 (1) d 10～d 23 のそれぞれを8台のプロセッサを使いソートする。ソートされた状態は図4の中段のようにキー値が降順に並べられたものになる。図中、三角形は各ブロックにおけるキー値の大きさの分布を示す。

(2) 別に一組の作業領域 e10～e13、d20～d23 を設けて図4の中段のように並列マージを実行する。その結果、下段のような結果が得られる。

15 (3) 次に2段目のマージソートを行う。（e 10、e 11）と（e 20、e 21）を組み合わせ、キー値の大小を考慮して2分割して2組のマージ入力用データ列の対を作り4並列のマージ操作を実行する（図5上段）。同時に、（e 12、e 13）と（e 22、e 23）を組み合わせ、キー値の大小を考慮して2分割して2組のマージ入力用データ列対を作り4並列のマージ操作を実行する（図5下段）。結局、4並列のマージが2系統並列に実行される。

20 (4) 次に3段目のマージソートを行う。d 10～d 13 と d 20～d 23 を組み合わせてキー値の大小を考慮して4分割して4組のマージ入力対を作り、e 10～e 23 へ出力する。各領域は4分割され、それぞれ正順・逆順の処理がされるので、8つの処理が並列に実行される。

25 図3の各段階におけるマージソート処理についてさらに詳しく説明する。図7はその説明図である。入力領域（D₁、n）と（D₂、n）の入力データ列の対が与えられたとする。ここで（D₁、n）と（D₂、n）はn個のデータからなるデータ列D₁とD₂を意味する。図7の各領域において左側ほどキー値が大きく、右側ほどキー値が小さくなるようにソートされている。

（段落0048） 図1に示すデータ列対分割部 2により（D₁、n）と（D₂、n）の整列済部分データ列対がキー値の大小を考慮して分割（図7では2分割）され、（D₁₁、n₁₁）と（D₂₁、n₂₁）の対と（D₁₂、n₁₂）と（D₂₂、n₂₂）の対とに分けられたとする。

この分割では下記の式を満足しなければならない。

30 D₁₁の末尾のデータのキー値とD₂₁の末尾のデータのキー値の小さい方

$\geq D_{12}$ の頭部のデータのキー値と D_{22} の頭部のデータのキー値の大きい方

$$n_{11} + n_{21} = 2x, \quad n_{12} + n_{22} = 2(n - x)$$

ここで $2x$ は、データ列対 $\{ (D_{11}, n_{11}), (D_{21}, n_{21}) \}$ がもつデータ数である。

マージ操作の入力であるデータ列対 (D_{11}, n_{11}) と (D_{21}, n_{21}) に対して、プロセッサ

5 P1 で正順併合を行うとともに、プロセッサ P2 で逆順併合を行う。

図中の点線は併合の順序を示す。その結果を出力領域 $(D_3, 2n)$ の部分である (D_{31}, n_{31}) に順次出力する。

同様に、整列済データ列 (D_{12}, n_{12}) と (D_{22}, n_{22}) の対に対して、プロセッサ P3 で正順併合を行うと共にプロセッサ P4 で逆順併合を行う。その結果を出力領域 $(D_3, 2n)$ の部分である (D_{32}, n_{32}) に順次出力する。

ここで、 $(D_{31}$ の末尾のデータのキー値) $\geq (D_{32}$ の頭部のデータのキー値) 、
 $n_{31} + n_{32} = 2n$ である。

以上のように入力データ列対を分割すると共に、分割されたデータ列対それぞれにプロセッサを割り当てるにより、複数のプロセッサによる同時並行処理が可能になる。なお、以上の説明はあくまで一例であり、本発明はこれに限定されない。

次に整列済データ列対への分割手順について説明する。図 7 の例からも明らかのように、マージソートが正しく行われるためには、 $(D_{31}$ の末尾のデータのキー値) $\geq (D_{32}$ の頭部のデータのキー値) でなければならない。 D_{31} は D_{11} 、 D_{21} のマージソート結果であり、 D_{32} は D_{12} 、 D_{22} のマージソート結果である。したがって、 $(D_{11}$ または D_{21} に含まれるデータのキー値の最小値) $\geq (D_{12}$ または D_{22} に含まれるデータのキー値の最大値) ということである。このことから本発明の領域分割の条件は次のようなものである。

前記分割条件では、予め整列された 2 つのデータ列の対 (D_1, n) と (D_2, n) が、部分データ列 (D_{11}, n_{11}) と (D_{21}, n_{21}) の対と、

部分データ列 (D_{12}, n_{12}) と (D_{22}, n_{22}) の対に 2 分割されたとき、

25 (部分データ列 D_{11} と D_{21} の末尾のデータのキー値の小さい方)

\geq (部分データ列 D_{12} と D_{22} の頭部のデータのキー値の大きい方)

が成立すると共に、

$$n_{11} + n_{21} = 2x, \quad n_{12} + n_{22} = 2n - 2x$$

が成立しなければならない。ここで x は D_1 と D_2 の頭部から数えたデータ列対のデータ数の半

30 分の数である。

例えば、1000個の整列済データ列対に対し、 $x = 300$ と $x = 600$ の2回の分割操作をすると、各データ列対の合計データ数が600、600、800の3つのデータ列対 { (D_{11}, n_{11}) , (D_{21}, n_{21}) } { (D_{12}, n_{12}) , (D_{22}, n_{22}) } { (D_{13}, n_{13}) , (D_{23}, n_{23}) } に分割される。この場合、両データ列を合わせ、先頭から600番目のデータのキー値 (5 D_{11} と D_{21} の末尾のデータキーの小さい方) と601番目のデータのキー値 (D_{12} と D_{22} の頭部のキーの大きい方)、1200番目のデータのキー値と1201番目のデータのキー値に前記の大小関係が満たされねばならない。

以上の条件を満たす整列済データ列対分割の手法を説明する。以下の正順併合用の分割手法または逆順併合用の分割手法、あるいは両方の手法を用いることにより、上記条件を満たすデータ列対への分割が実現できる。

整列済データ列対の分割手法の具体的な手順の説明の前に、前提となる事項について説明する。それぞれ n 個のデータからなる 2 つの整列済データ列 D_1 と D_2 を仮定し、両データ列を (n/k) 組のデータを含む k 組のデータ列対に均等に分割する。

個々のデータ列は 0 から $k-1$ のデータ列番号で指定する。また各データの位置は、データ列頭部を 0 としデータ毎に 1 ずつ増加する指標値で指定する。この位置指定法は D_1 、 D_2 全体のデータ位置指定のほか、分割した各データ列内の位置指定にも用いる。また、データ列の頭部あるいは尾部という表現で始端あるいは終端のデータ位置を指定する。

前述した均等分割後の D_1 と D_2 の部分データ列中、同一データ列番号のデータ列同志を仮の併合対象データ列対とする。

20 すなわち、データ列番号 i ($i=0 \sim k-1$) のデータ列対を対象とする正順併合の始点は、 D_1 と D_2 の始端から $(n/k)(i)$ 組のデータを除いた残りのデータ列の始端、すなわち指標値が $(n/k)(i)$ の点とする。

25 またデータ列番号 i ($i=0 \sim k-1$) のデータ列を対象とする逆順併合の始点は、 D_1 と D_2 の始端からそれぞれ $(n/k)(i+1)$ 組のデータが存在するデータ列の末尾、換言すれば指標値が $(n/k)(i+1)-1$ の点とする。

D_1 と D_2 内のキーの分布は互いに独立である。このため、それぞれのデータ列内で整列済であっても、仮の開始点から併合操作をしたのでは、全体としてキーの分布に矛盾のない整列結果を得ることはできない。そこで以下の手順により前記条件を満たす開始点（分割点）を求める。

<正順併合の場合>

それぞれ整列された n 個のデータからなる 2 つのデータ列、 D_1 と D_2 とがある。段落 0048 の条件を満たし、かつ、両データ列の先頭からのキー値の大小を考慮してデータ列の先頭からのデータ数が $2(n/k)+2$ 、 $4(n/k)+2$ 、 $6(n/k)+2$ 、 \dots $2(k-1)(n/k)+2$ となるデータ列対を識別する。この境界はデータ数が $2(n/k)$ 、 $4(n/k)$ 、 $6(n/k)$ 、 \dots $2(k-1)(n/k)$ のデータ列対を除いた残り

5 のデータ列対の頭部の指標である。

図 8 を参照する。

S10：境界初期値（あるいは仮の境界値）として (n/k) 、 $2(n/k)$ 、 $3(n/k)$ 、 \dots $(k-1)(n/k)$ の指標値をもつ点を選ぶ。

S11： D_1 と D_2 域の作業用指標変数 i と j とに、S10 で定めた境界初期値の指標値を設定する。

S12：下記の手順を大小関係が反転する迄繰り返す。ただし最初の比較を含め、キー値が等しい場合も反転とみなす。

S12a： $D_1(i)$ と $D_2(j)$ のキー値を比較する。

S12b：大小関係を判別する。

15 (1) 最初の比較で両者が等しいなら S13a に分岐する。

(2) 大小関係が逆転したら S13b に分岐する。

(3) 最初の比較で両者が等しくない場合、または 2 回目以降の比較で大小関係が反転しない場合、キー値の大きい方の指標値に 1 を加え、小さい方の指標値から 1 を引いた後、S12 に分岐する (S12c)。

20 S13a： i と j を境界点の指標とし処理を終わる。

S13b：大小関係が変化する直前と直後の比較操作の対象データ中、大きいキー同志を比較し、その小さい方を境界のデータとする。他方の領域の境界データには今求めたデータの当初の比較相手を選ぶ。

<逆順併合の場合>

25 それぞれ n 個のデータを収容した 2 つのデータ列、 D_1 と D_2 とがある。段落 0048 の条件を満たし、かつ、両データ列のキー値の大小分布を考慮してデータ列先頭からのデータ数が $2(n/k)$ 、 $4(n/k)$ 、 $6(n/k)$ 、 \dots $2(k-1)(n/k)$ となるデータ列の末尾を求める。以下では図 9 を参照する。

S20：境界初期値（仮の境界値）として、それぞれ $(n/k)-1$ 、 $2(n/k)-1$ 、 $3(n/k)-1$ 、 \dots $(k-1)(n/k)-1$ の指標値をもつ点を選ぶ。

30 S21： D_1 と D_2 域の作業用指標変数 i と j とに S20 で定めた境界初期値の指標値を設定する。

S 2 2 : 下記の手順を大小関係が反転する迄繰り返す。ただし最初の比較を含め、キー値が等しい場合も反転とみなす。

S 2 2 a : $D_1(i)$ と $D_2(j)$ のキー値を比較する。

S 2 2 b : 大小関係を判別する。

- 5 (1) 最初の比較で両者が等しいなら S 2 3 a に分岐する。
 - (2) 大小関係が反転したら S 2 3 b に分岐する。
 - (3) 最初の比較で両者が等しくない場合、または 2 回目以降の比較で大小関係が反転しない場合、キー値の大きい方の指標値に 1 を加え、小さい方の指標値から 1 を引いて S 2 2 に分岐する。
- S 2 3 a : i と j を境界点の指標とし処理を終わる。
- 10 S 2 3 b : 大小関係が変化する直前と直後の比較操作の対象データ中、小さいキー値同志を比較し、その大きい方を境界のデータとする。また、他方の領域の境界のデータには上で求めたデータの当初の比較相手を選ぶ。

上記手順によるデータ列対の分割例について説明する。

図 1 0 は降順に整列された、それぞれ 10 個の数値を含む下記の 2 つのデータ列を示す。図 1 0

- 15 (a) は正順併合の始点の決定手順の説明図、図 1 0 (b) は逆順併合の始点の決定手順の説明図、図 1 0 (c) は領域の分割結果を示す。

D_1 : (1948, 1868, 1862, 1148, 740, 620, 588, 256, 112, 100)

D_2 : (1996, 1992, 1966, 1694, 1544, 1448, 1182, 432, 254, 30)

(段落 0 0 5 9) <正順併合の始点の決定>

- 20 (1) データ列頭部から 5 組のデータを除いたデータ列の頭部が仮の分割点となる。作業用指標 i , j にこの値を設定する。 $i=j=5$ となる。
- (2) $D_1(i=5)=620$ と $D_2(j=5)=1466$ を比較する。1466 が大きいので、 $i=i-1=4$, $j=j+1=6$ とする。
- (3) $D_1(i=4)=740$ と $D_2(j=6)=1182$ を比較する。1182 が大きいので、 $i=i-1=3$, $j=j+1=7$ とする。
- (4) $D_1(i=3)=1148$ と $D_2(j=7)=432$ を比較する。1148 が大きいので、ここで大小関係の反転を
- 25 認識する。
- (1) 反転前後の比較で値の大きい 1148 と 1182 を比較し、その小さい方である 1148 の点を一方の境界とする。他方の境界は 1148 の当初の比較相手である 432 を得る。なお、一方の境界が判明すれば、次式(1)を用い $2 \times 5 - 3 = 7$ として他方の境界指標値を求めることもできる。
- 他方の開始点の指標値 = 仮併合開始点の指標 $\times 2$ - 既知の開始点の指標
- 30 <逆順併合の始点の決定>

(1) データ列頭部から 5 組のデータがある点が仮の分割点となる。作業用指標 i, j にこの値を設定する。 $i=j=4$ とする。

(2) $D_1(i=4)=740$ と $D_2(j=4)=1544$ とを比較すると、1544 が大きいので、 $i=i-1=3, j=j+1=5$ とする。

5 (3) $D_1(i=3)=1148$ と $D_2(j=5)=1446$ とを比較すると、1446 が大きいので $i=i-1=2, j=j+1=6$ とする。

(4) $D_1(i=2)=1862$ と $D_2(j=6)=1182$ とを比較すると、1862 が大きいので大小関係が反転したことになる。

(5) 反転前後の比較で値の小さい 1148 と 1182 を比較し、大きい方の 1182 を一方の境界とする。

10 他方の境界は 1182 の当初の比較相手のデータ 1862 を得る。段落 0 0 5 9 の式を用い、 $2 \times 4 - 6 = 2$ として他方の境界の指標値を求めることもできる。

他の例を図 11 (a) (b) に示す。図 11 (a) は、 $n = 12, k = 4, n/k = 3$ の場合である。図 11 (b) は、 $n = 10, k = 5, n/k = 2$ の場合である。この例のように、分割後の部分データ列により併合する場合の出力データが D_1 または D_2 の一方にのみ存在する場合(両

15 方のデータ列にまたがっていない場合) であっても、前記条件を満たすことがわかる。

本発明の実施の形態に係る装置／方法において、分割する区画の数は使用可能な処理装置の数に応じて任意に決めることができる。

本発明に係る装置／方法において、マージソートの入力となる整列済部分データ列対を複数のデータ列対に分割する時間の全時間に占める割合を小さくするため、データ数が大きいほど効果 20 も大きい。なお、複数のデータ列対に分割する処理のうち、実質的な併合開始点を決定する部分をマージソート作業を割り当てられたプロセッサ自身が実行すれば、より充実した並列処理が可能となり全処理時間に占める領域対分割にかかる時間の割合を削減できる。

(1) 効果の大きい場合

データ列 D_1 と D_2 のキー値の分布がほぼ同じ場合は短時間に大小関係が反転する。このためマ 25 ージ入力となるデータ境界の決定に多くの時間を要しない。

(2) 分割に時間のかかる場合

下記のケースが考えられる。

(a) 分割対の確定に時間のかかる場合

k 区画に分割する場合で対象となるデータ列対の区画番号が $k/2$ に近い場合で、 D_1 または $D_{30} D_2$ の頭部や尾部、換言すればデータ列の端部近辺になってようやく大小関係が反転する場合、大小

関係の反転までに時間がかかる。キーの分布が D_1 と D_2 とで偏りがある場合に発生するが、このようなケースが頻発することは少ないと考える。

(b) キー値の価域に重複がない場合と大小関係が反転しない場合

大小関係の反転を求めて比較を続け、領域の頭部または端部に至っても反転しない場合である。

- 5 このときは次のように扱うことができる。すなわち、頭部に達しても反転しない場合は頭部の外側に無限大のキーを仮定して強制的に反転させ、尾部に至っても反転しない場合は尾部の外に無限小のキーを仮定して反転させる。

なお、上記例ではデータ列 D_1 と D_2 のデータ数は同数を仮定していた。本発明はこれに限定されない。整列タスクや併合タスクにおいてデータ数を指定しさえすれば D_1 と D_2 のデータの数が異なる場合にも適用できる。

本発明は汎用的な技術であり、主記憶を共有する汎用的な並列プロセッサ方式の電子計算機に特に適合性が高い。しかし整列対象のデータ群が共通にアクセスできる環境にあればシステムの構成や処理形態を問わず適用できると考えられる。本発明は、並列プロセッサ方式の電子計算機のソフトウェア（OS）に組み込まれるプログラムのひとつとして実現されると考えられる。

15

発明の実施の形態 2.

発明の実施の形態 1 の装置／方法は、図 2 に示すように正順併合と逆順併合を同時に用いるものであった。この発明は正順併合と逆順併合のいずれか一方しか用いない場合にも適用できる。

- 20 図 1 2 は並列 2 分木マージソート法に適用した場合のフローチャートである。この場合は初段のマージから 2 分割するので常に p 個のプロセッサが動作する。マージ処理を行う並列プロセッサも正順併合か逆順併合のいずれかの機能を装備するだけでよいので実現は簡単になる。

実施形態 2 の処理内容は発明の実施の形態 1 の場合と同様であるので、その説明を省略する。

なお、図 1 2 が図 2 と異なる部分は S 4 b と S 5 b と S 6 b である。

- 25 本発明は、以上の実施の形態に限定されることなく、特許請求の範囲に記載された発明の範囲内で、種々の変更が可能であり、それらも本発明の範囲内に包含されるものであることは言うまでもない。

- 30 また、本明細書において、手段とは必ずしも物理的手段を意味するものではなく、各手段の機能が、ソフトウェアによって実現される場合も包含する。さらに、一つの手段の機能が、二つ以上の物理的手段により実現されても、若しくは、二つ以上の手段の機能が、一つの物理的手段により実現されてもよい。

請求の範囲

1. 複数のプロセッサを含み、2つの入力データ列に対してマージソート処理を行うマージソート用並列プロセッサと、マージソート処理の入力となる2つの整列済データ列対（つい）を5指定された組数の部分データ列対に細分割するデータ列対分割用並列プロセッサのほか、全体を制御する管理用プロセッサを備え、前記管理用プロセッサは前記分割用並列プロセッサ群から受けた複数の部分データ列対を前記マージソート用並列プロセッサにそれぞれ割り当て、並列にマージ処理を行わせることを特徴とする並列マージソート処理装置。
- 10 2. 部分データ列の対を割り当てられた前記マージソート用並列プロセッサはそれぞれ独立にマージ処理を行いその結果を予め指示された出力域に出力する。これらの出力域の全体が最終マージ結果となるか、あるいは中間マージ結果として次段の並列マージに使われるように出力されることを特徴とする請求項1記載の並列マージソート処理装置。
- 15 3. 前記部分データ列対それぞれに対して2つのプロセッサが割り当てられ、第1のプロセッサは正順併合（前記部分データ列対のキー値の大きい側の端から小さい方に向かってマージ処理を行い、前記出力領域のキー値の大きい側の端から小さい方に向かって順に出力を書き込む）を行い、第2のプロセッサは逆順併合（前記部分データ列対のキー値の小さい側の端から大きい方に向かってマージ処理を行い、前記出力領域のキー値の小さい側の端から大きい方に向かって順に出力を書き込む）を行うことを特徴とする請求項2記載の並列マージソート処理装置。20
4. 前記分割条件として、下記の項目を満たすことを特徴とする請求項1記載の並列マージソート処理装置。
 - ◆以下ではn個のデータからなる整列済データ列Dを(D, n)で表し、2つのデータ列を対にしたものを作成する。したものを{(D₁, n), (D₂, n)}と表記する。
 - ◆整列済データ列対：{(D₁, n), (D₂, n)}が、部分データ列対：{(D₁₁, n₁₁), (D₂₁, n₂₁)}と{(D₁₂, n₁₂), (D₂₂, n₂₂)}に2分割されたとき、(部分データ列D₁₁とD₂₁の末尾のデータのキー値の小さい方)
≥ (部分データ列D₁₂とD₂₂の頭部のデータのキー値の大きい方)
- 30 が成立すると共に、

$$n_{11} + n_{21} = 2x, \quad n_{12} + n_{22} = 2(n - x)$$

が成立する。ここで x は D_1 と D_2 の頭部から数えた領域対のデータ数の半分である。

5. 並列プロセッサを用いて並列マージソート処理を行うための方法であって、

5 入力として、未整列の 1 つのデータ列、または 2 つの整列済部分データ列、あるいは 3 つ以上の整列済部分データ列のいずれかとその属性情報（種別・データの所在情報や個数など）を受け入れ、

整列対象が未整列データ列の場合はそのデータ列を分割して複数の未整列部分データ列を得る第 1 ステップと、

10 前記複数の部分データ列にそれぞれプロセッサを割り当てる第 2 ステップと、

割り当てられたプロセッサにより前記複数の部分データ列をそれぞれ独立に任意のアルゴリズム例えればクイックソート法で整列し、整列された部分データ列を得る第 3 ステップと、

第 3 ステップまたは第 7 ステップで得たか初期データとして入力された、任意の 2 つの整列済部分データ列でマージソート用入力データ列の対をつくり、その対を所定の分割条件のもとでそれぞれ要求された組数の部分データ列の対に分割する第 4 ステップと、

分割された複数の部分データ列対から並列マージソート用のジョブ情報を編集する第 5 ステップと、

得られた並列マージソート用ジョブにマージソート用プロセッサ群を割り当てる第 6 ステップと、

20 割り当てられたプロセッサによりマージ処理を行う第 7 ステップと、

マージ処理された整列済部分データ列を前記整列済部分データ列として前記第 4 ステップ乃至前記第 7 ステップを繰り返すステップとを備える並列マージソート処理方法。

但し、以上は 1 つの未整列データ列が与えられる場合に適用されるもので、2 つ以上の整列済データ列が与えられる場合は第 1 ステップから第 3 ステップ迄は不要である。

25 また第 4 ステップ乃至前記第 7 ステップの繰り返しはマージ処理された整列済部分データ列が 1 つのデータ列になった段階で終る。特に 2 つの整列済データ列が与えられる場合は第 4 ステップから第 7 ステップは 1 回だけ行われ繰り返しは必要としない。

6. 前記第 6 ステップにおいて前記部分データ列の対に対して 2 つのプロセッサが割り当て

30 られ、前記第 7 ステップにおいて第 1 のプロセッサは前記部分データ列のキー値の大きい側から

の正順併合処理を、第2のプロセッサは同じ部分データ列のキー値の小さい側の端からの逆順併合処理を、それぞれ同時に行われるように第5ステップでジョブ制御情報を編集することを特徴とする請求項5記載の並列マージソート処理方法。

- 5 7. 前記第4ステップの分割条件として下記の項目を満たす機能をもつことを特徴とする請求項5記載の並列マージソート処理方法。

データ列の対 (D_1, n) と (D_2, n) が、部分データ列対 $\{ (D_{11}, n_{11}), (D_{21}, n_{21}) \}$ と、部分データ列： $\{ (D_{12}, n_{12}), (D_{22}, n_{22}) \}$ に2分割されたとき、
(部分データ列 D_{11} と D_{21} の末尾のデータのキー値の小さい方)

10 \geq (部分データ列 D_{12} と D_{22} の頭部のデータのキー値の大きい方)

が成立すると共に、 $n_{11} + n_{21} = 2x$, $n_{12} + n_{22} = 2(n - x)$ が成立する。

ここで x は部分データ列対 $\{ (D_{11}, n_{11}), (D_{21}, n_{21}) \}$ のデータ数の半分の値であり、 $n_{11} = n_{21}$ の場合の D_{11} と D_{21} のデータ数でもある。

- 15 8. 前記第4ステップとして下記の機能を備えることを特徴とする請求項5記載の並列マージソート処理方法。

(1) 整列済データ列の対： $\{ (D_1, n), (D_2, n) \}$ を k 組の区画対に分割する操作は、両データ列のキー値の大小を考慮して D_1 と D_2 の頭部から数えたデータ数の合計が $2x$ となる2分割操作を、 x の値を変えて $k - 1$ 通り行うことと等価となる。

20 この場合、整列済データ列対： $\{ (D_1, n), (D_2, n) \}$ の k 組の区画対への細分割問題は、前述した請求項7を満たすデータ列の2分割問題に置き換えられる。

(2) データ列中のデータ位置は指標値で指定する。この値はデータ列 D_1 または D_2 の先頭データの指標値を0とし、順次1、2、…と増加する番号をつける。

また、請求項7の x はデータの個数を表すものだが、 x の値自体を指標値と見なすとデータ列25の先頭から1、2、3と数えて $x + 1$ 番目のデータを指すことになる。

さらに請求項7で $n_{11} = n_{21}$ の場合、データ数を表す式 $n_{11} = n_{21} = x$ は先頭から数えて x 番目のデータ、指標値で言えば $x - 1$ のデータ位置が D_1 と D_2 の分割境界であることを表していると解釈することもできる。

(3) 領域の分割機能

◆前記データ列 D_1 と D_2 内の個々のデータを指定する指標変数*i*と*j*に境界指標値の初期値として前記xを設定（例えば10000個ずつのデータ列対を1000個ずつのデータ列対10区画に分割する場合、個数としてのx=1000、分割位置の初期指標には1000, … 9000を設定）するステップと、

5 ◆データ列 D_1 の指標変数*i*が指すデータのキー値とデータ列 D_2 の指標変数*j*が指すデータのキー値を比較する比較ステップと、

◆最初の比較で D_1 の指標変数*i*が指すデータのキー値と D_2 の指標変数*j*が指すデータのキー値が等しくない場合、キー値の大きい方の指標変数に1を加え、小さい方の指標変数から1を引いた後、前記比較ステップに分岐するステップと、

10 ◆2回目以降の比較で D_1 の指標変数*i*が指すデータのキー値と D_2 の指標変数*j*が指すデータのキー値との大小関係に変化がない場合、大きい方の指標変数に1を加え、小さい方の指標変数から1を引いた後、前記比較ステップに分岐するステップと、

◆最初の比較で D_1 の指標変数*i*が指すデータのキー値と D_2 の指標変数*j*が指すデータのキー値とが等しいとき、指標変数*i*が指すデータと指標変数*j*が指すデータをそれぞれ分割境界とするステップと、

◆ D_1 の指標変数*i*が指すデータのキー値と D_2 の指標変数*j*が指すデータのキー値との間の大小関係が前回の大小関係から反転（2回目以降の比較では等しい場合も反転とみなす）したとき、前回の比較操作における D_1 のキー値と D_2 のキー値のうちの大きい方と、今回の比較操作における D_1 のキー値と D_2 のキー値の内の大きい方とを比較して、小さい方のキー値をもつデータを分割境界とすると共に、当該データの当初の比較相手を他方の境界とするステップ

20 とを備えることを特徴とする請求項5記載の並列マージソート処理方法。これは2分した下位データ列対の頭部の位置の決定に用いる。

9. 前記第4ステップとして下記の機能を備えることを特徴とする請求項5記載の並列マージソート処理方法。

(1) 整列済データ列の対：{(D_1 , n), (D_2 , n)}をk組の区画対に分割する操作は、両データ列のキー値の大小を考慮して D_1 と D_2 の頭部から数えたデータ数の合計が $2x$ となる2分割操作を、 x の値を変えて $k-1$ 通り行うことと等価となる。

30 このように見なすと、整列済データ列の対：{(D_1 , n), (D_2 , n)}のk組の区画対への細分割問題は、前述した請求項7の条件を満たすデータ列の2分割問題に置き換えられる。

(2) データ列中のデータ位置は指標値で指定する。この値は D_1 または D_2 の先頭データの指標値を 0 とし、順次 1、2、…と増加する。

また請求項 7 の x はデータの個数だが、 x の値自体を指標値と見なすとデータ列の先頭から 1, 2, 3 と数えて $x + 1$ 番目のデータを指すことになる。

5 さらに請求項 7 で $n_{11} = n_{21}$ の場合、 $n_{11} = n_{21} = x$ は先頭から数えて x 番目のデータ、指標値で言えば $x - 1$ のデータ位置に D_1 と D_2 の分割境界があることを表している。

(3) 領域の分割機能

◆前記データ列 D_1 と D_2 の個々のデータを指定する指標変数 i と j とに境界の指標値の初期値として前記 $x - 1$ を設定（例えば 1 0 0 0 0 個ずつのデータ列対を 1 0 0 0 個ずつのデータ列対 10 区画に分割する場合、個数としての x には 1 0 0 0、分割位置の初期指標には 9 9 9, 1 9 9 9, … 8 9 9 9 を設定）するステップと、

データ列 D_1 の指標変数 i が指すデータのキー値とデータ列 D_2 の指標変数 j が指すデータのキー値を比較する比較ステップと、

◆最初の比較で D_1 の指標変数 i が指すデータのキー値と D_2 の指標変数 j が指すデータのキー値が等しくない場合、キー値の大きい方の指標変数に 1 を加え、小さい方の指標変数から 1 を引いた後、前記比較ステップに分岐するステップと、

◆2 回目以降の比較で D_1 の指標変数 i が指すデータのキー値と D_2 の指標変数 j が指すデータのキー値との大小関係に変化がない場合、大きい方の指標変数に 1 を加え、小さい方の指標変数から 1 を引いた後、前記比較ステップに分岐するステップと、

20 ◆最初の比較で D_1 の指標変数 i が指すデータのキー値と D_2 の指標変数 j が指すデータのキー値とが等しいとき、指標変数 i が指すデータと指標 j が指すデータをそれぞれ分割境界とするステップと、

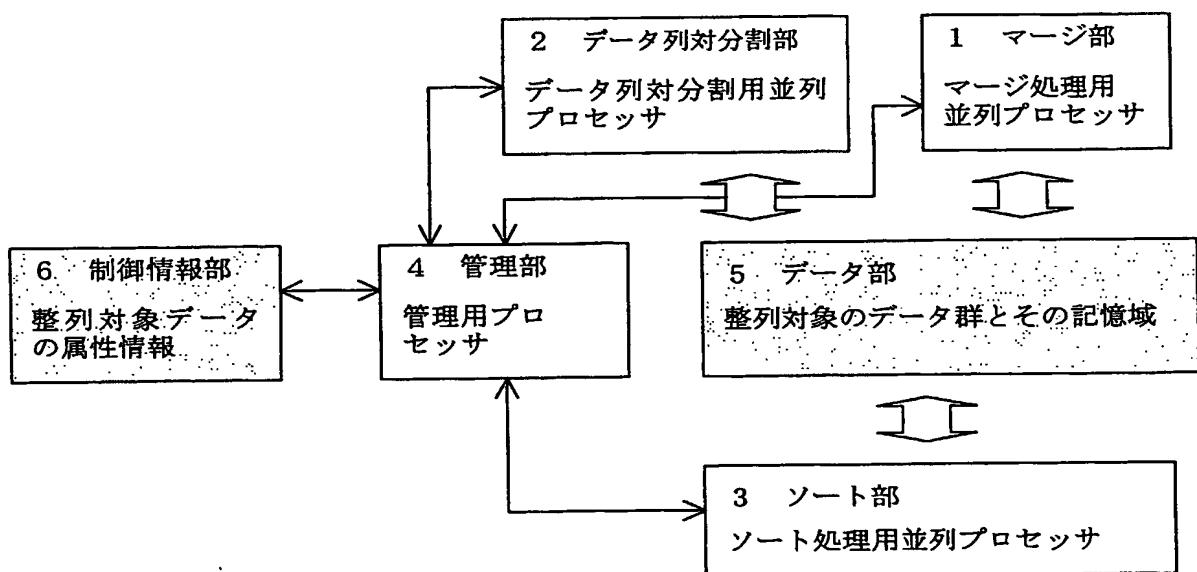
◆ D_1 の指標変数 i が指すデータのキー値と領域 D_2 の指標変数 j が指すデータのキー値との間の大小関係が前回の大小関係から逆転（2 回目の比較以降では等しい場合も反転とみなす）したとき、前回の比較操作における D_1 のキー値と D_2 のキー値のうちの小さい方と、今回の比較操作における D_1 のキー値と D_2 のキー値の内の小さい方とを比較して、大きい方のキー値をもつデータを分割境界とすると共に、当該データの当初の比較相手を他方の境界とするステップとを備えることを特徴とする請求項 5 記載の並列マージソート処理方法。これは 2 分した上位データ列対の末尾の決定に用いる。

10. 2組の整列済データ列の対を複数の整列済データ列対に細分割して任意の並列度のマージソート操作で整列させる並列マージソート処理方法であって、
整列された第1のデータ列及び第2のデータ列の対を用意するステップと、
要求される並列度に応じて前記第1のデータ列及び第2のデータ列の対を複数の整列済部分データ列対に細分割する分割ステップと、
5 細分割された整列済部分データ列の対を並列にマージするステップとを備え、
細分割した整列済部分データ列の対数に関係なく全体としてキーの並びに矛盾の無いデータ列の出力が可能な並列マージソート処理方法。
- 10 11. 複数のプロセッサを含む並列プロセッサを用いて並列マージソート処理を行うためのプログラムであって、
入力として、未整列の1つのデータ列、または2つの整列済部分データ列、あるいは3つ以上の整列済部分データ列のいずれかとその属性情報（種別・データの所在情報や個数など）を受け入れ、
15 整列対象が未整列データ列の場合はそのデータ列を分割して複数の未整列部分データ列を得る第1ステップと、
前記複数の部分データ列にそれぞれプロセッサを割り当てる第2ステップと、
割り当てられたプロセッサにより前記複数の部分データ列についてそれぞれ独立に任意のアルゴリズム、例えばクイックソート法で整列処理を行う第3ステップと、
20 第3ステップまたは第7ステップで得たか初期データとして入力された、任意の2つの整列済部分データ列でマージソート用の入力データ列の対をつくり、所定の分割条件のもとでそれぞれ要求された分割数の部分データ列対に細分割する第4ステップと、
分割された部分データ列対をマージソートするためのジョブ情報を編集する第5ステップと、
得られたマージソートジョブに対してプロセッサ群を割り当てる第6ステップと、
25 割り当てられたプロセッサにより並列にマージ処理を行う第7ステップと、
マージ処理されたデータ列を前記部分データ列として前記第4ステップ乃至前記第7ステップを繰り返すステップと、を実行させるためのプログラム。
但し、以上は1つの未整列データ列が与えられる場合に適用されるもので、2つ以上の整列済データ列が与えられる場合は第1ステップから第3ステップ迄は不要である。

また第4ステップ乃至前記第7ステップの繰り返しはマージ処理された整列済部分データ列が1つのデータ列になった段階で終る。特に2つの整列済データ列が与えられる場合は第4ステップから第7ステップは1回だけ行われ繰り返しは必要としない。

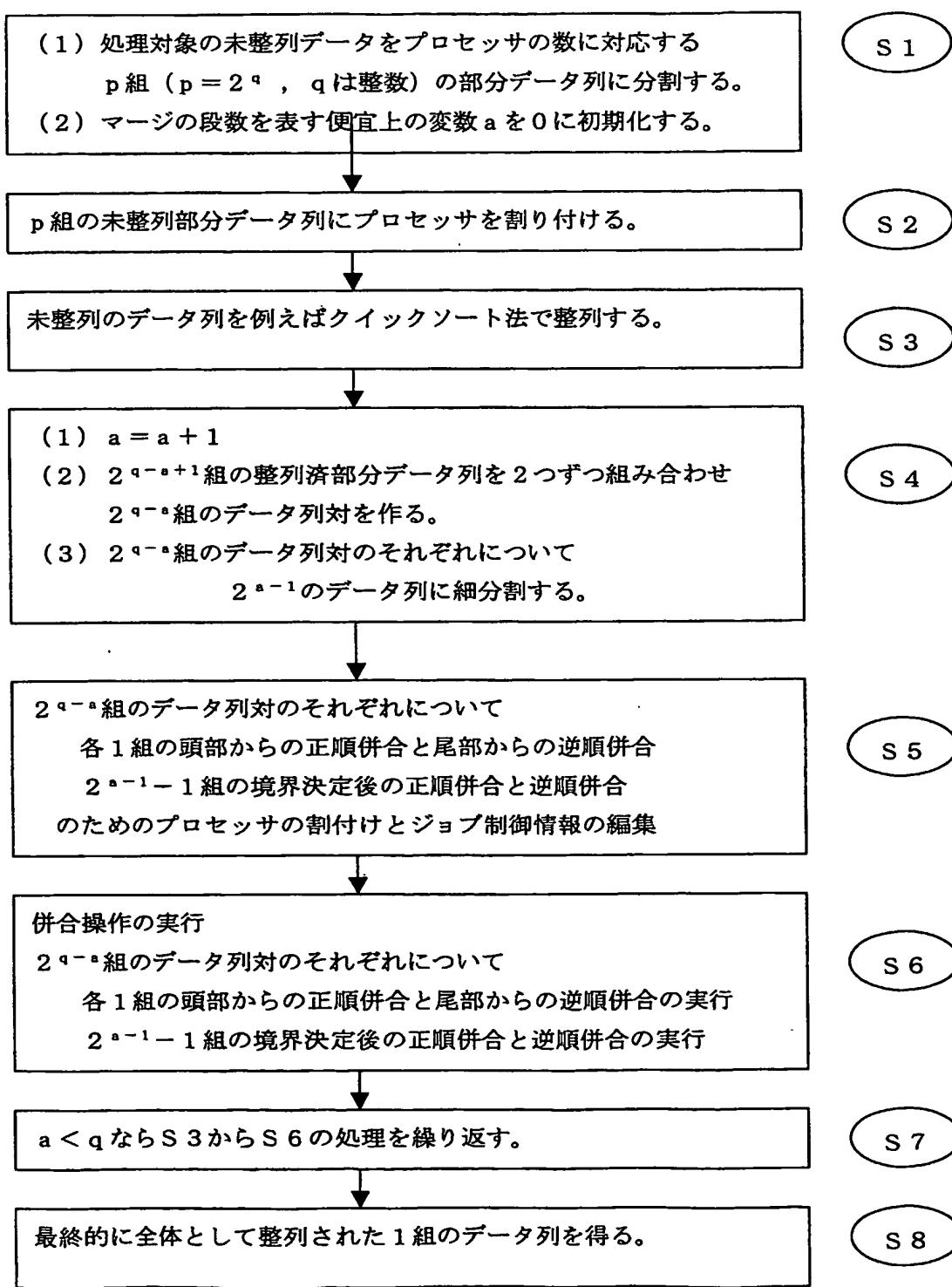
1 / 14

図 1



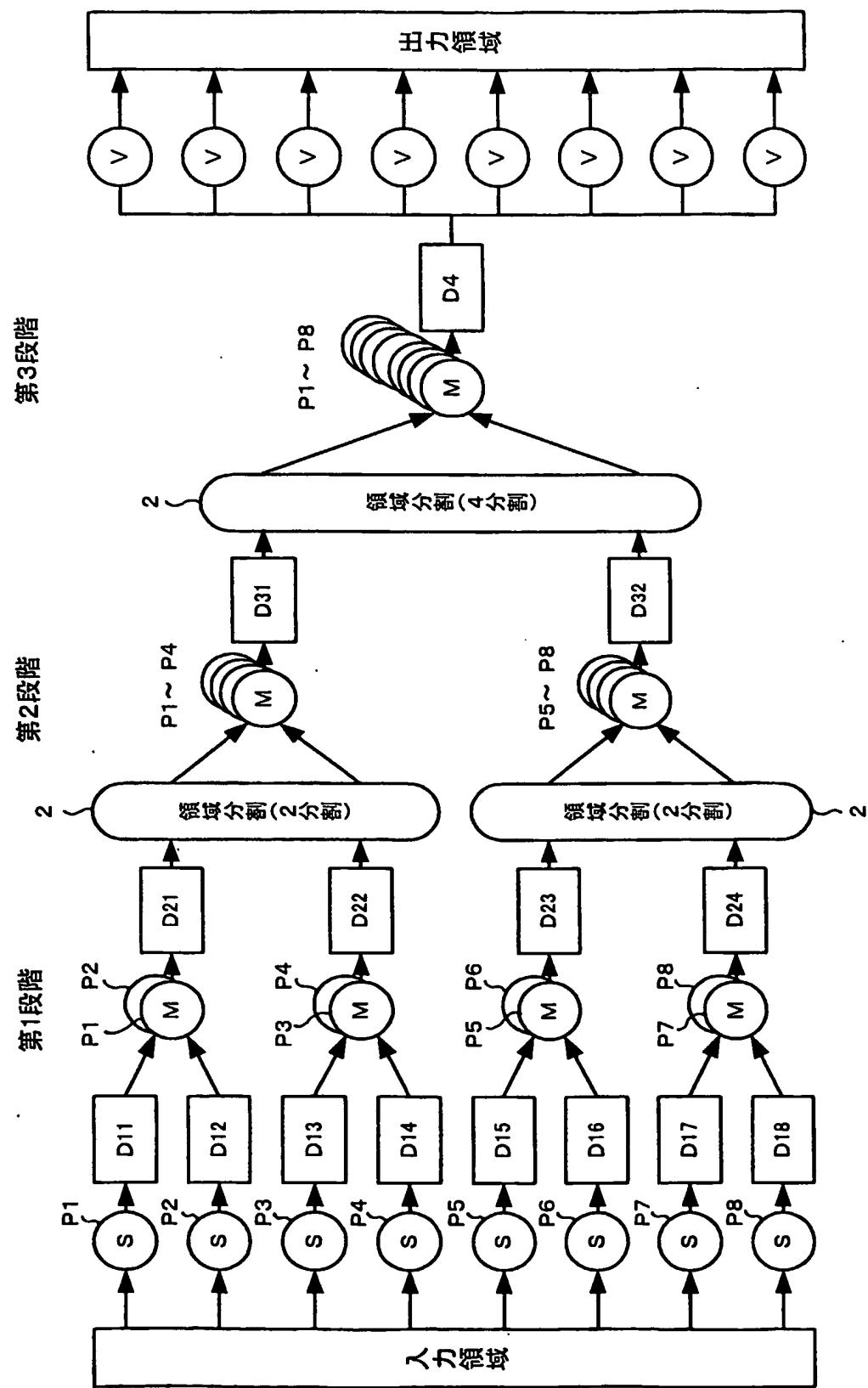
2 / 14

図 2



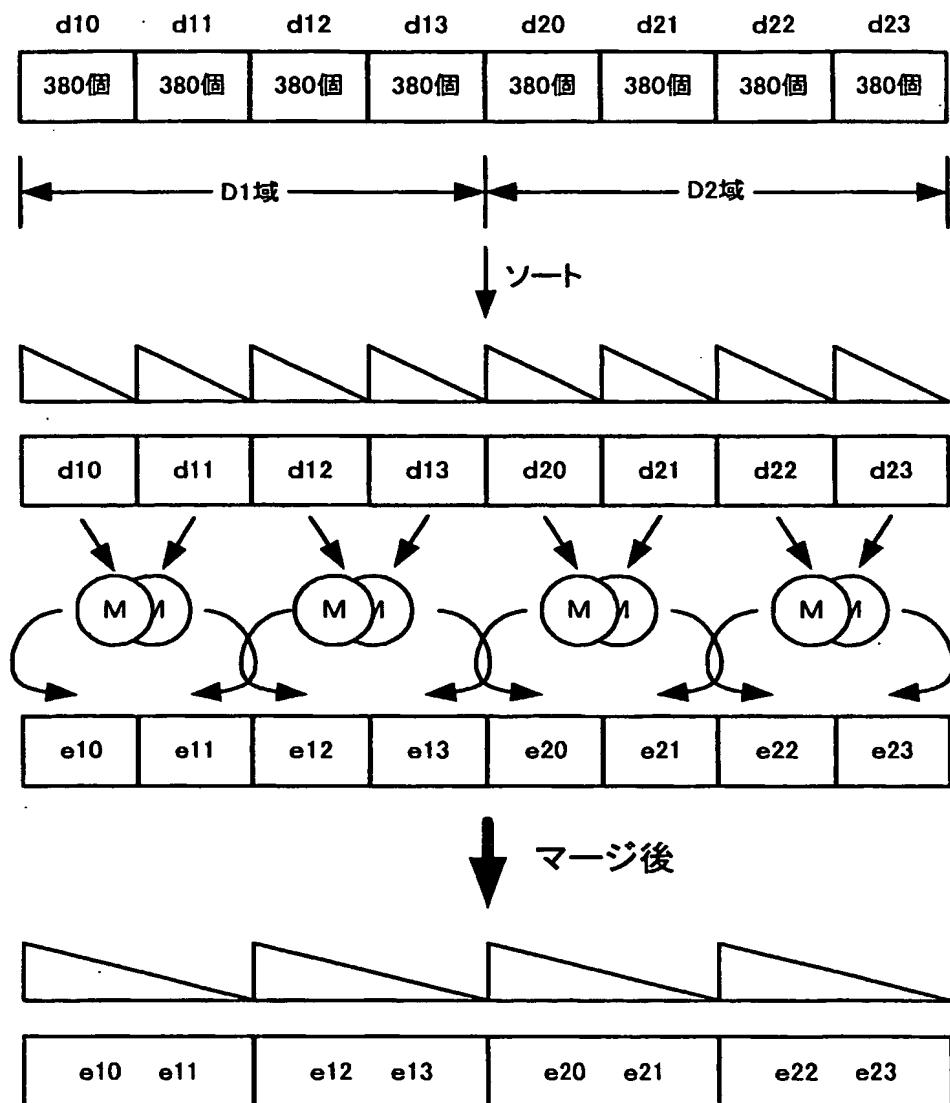
3 / 14

図 3



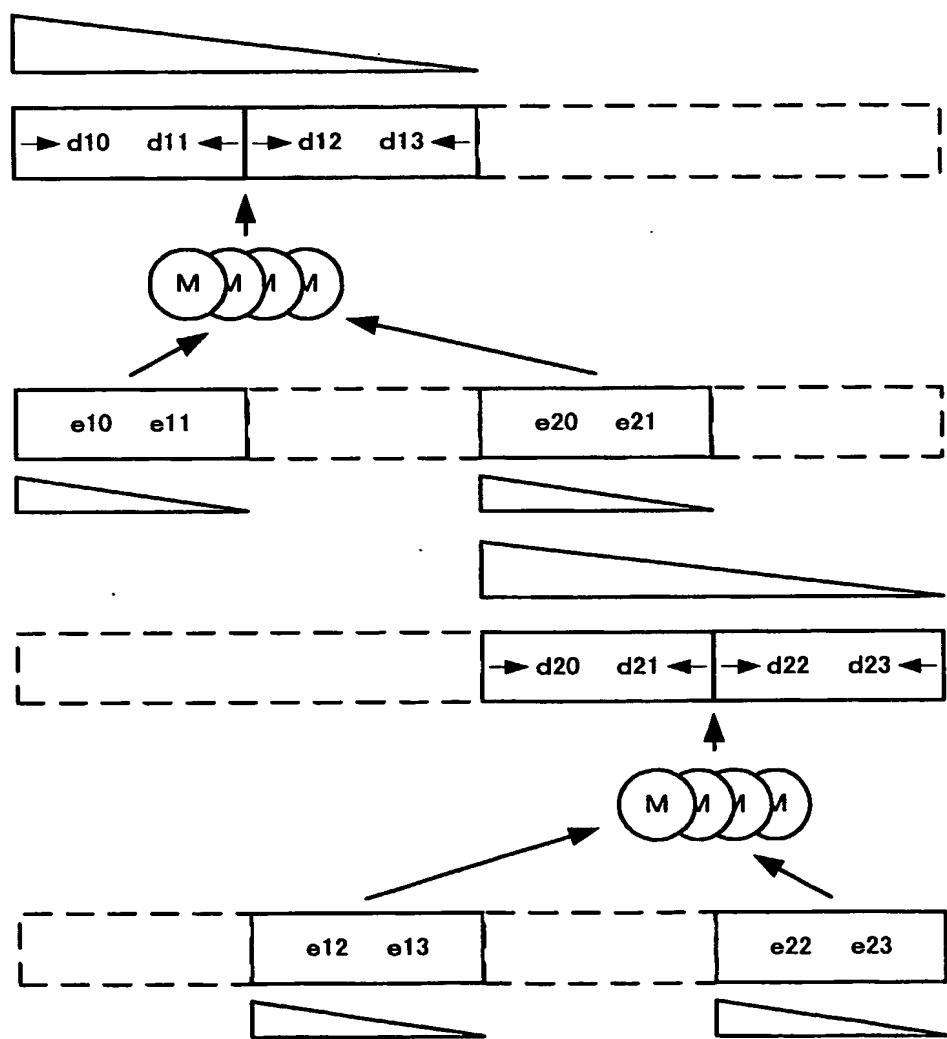
4 / 1 4

図 4



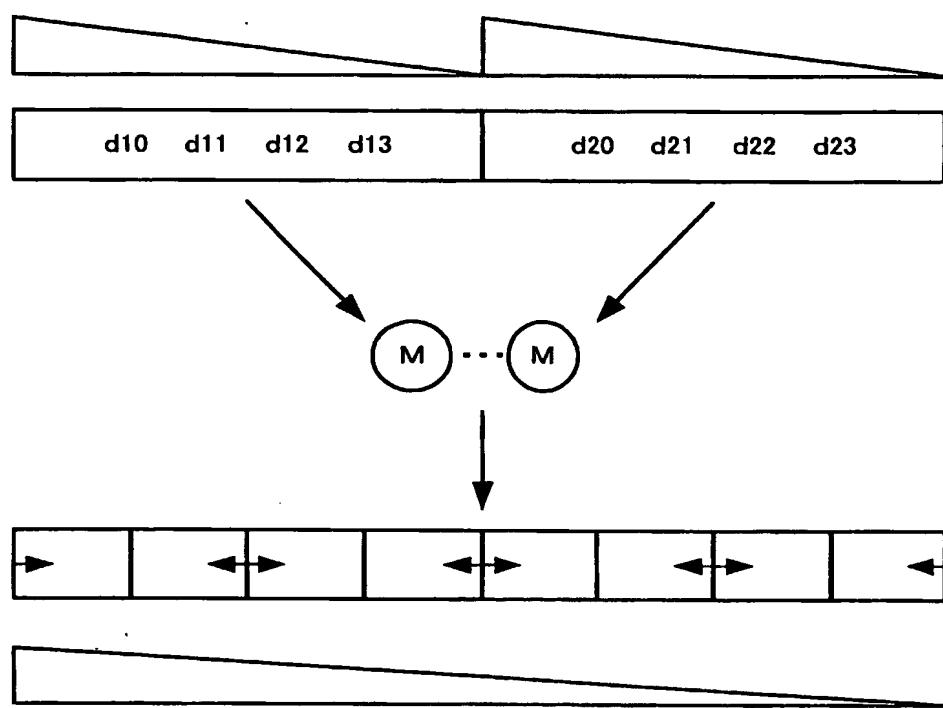
5 / 1 4

図 5



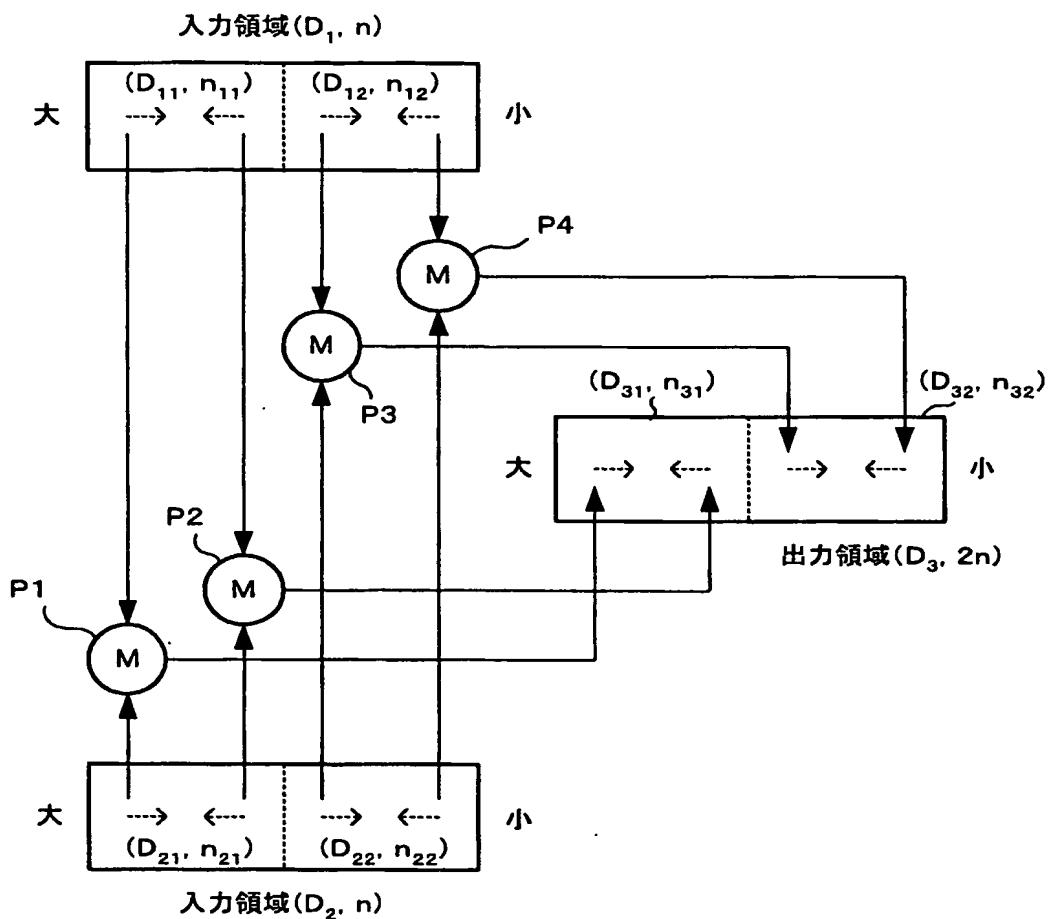
6 / 1 4

図 6



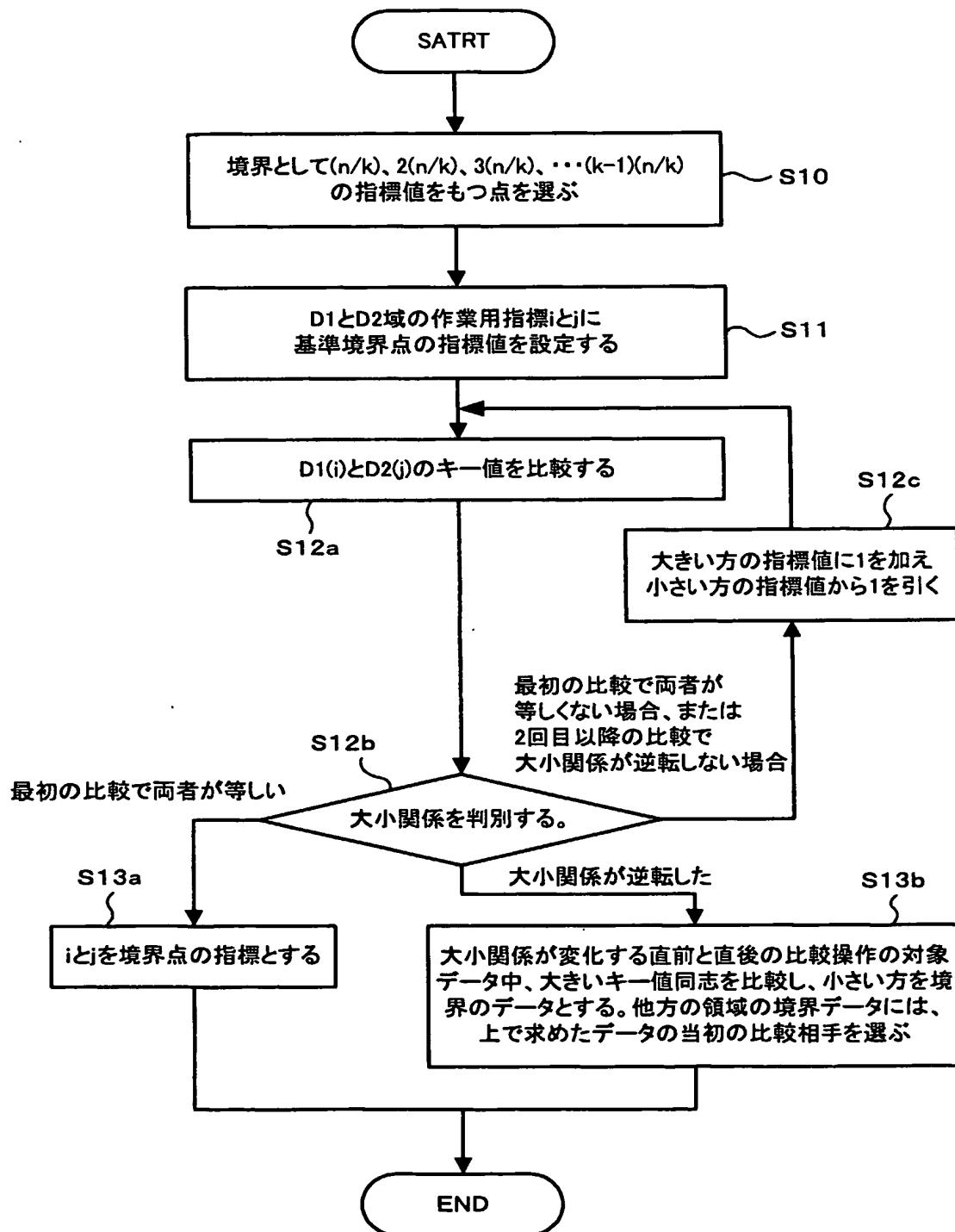
7 / 14

図 7



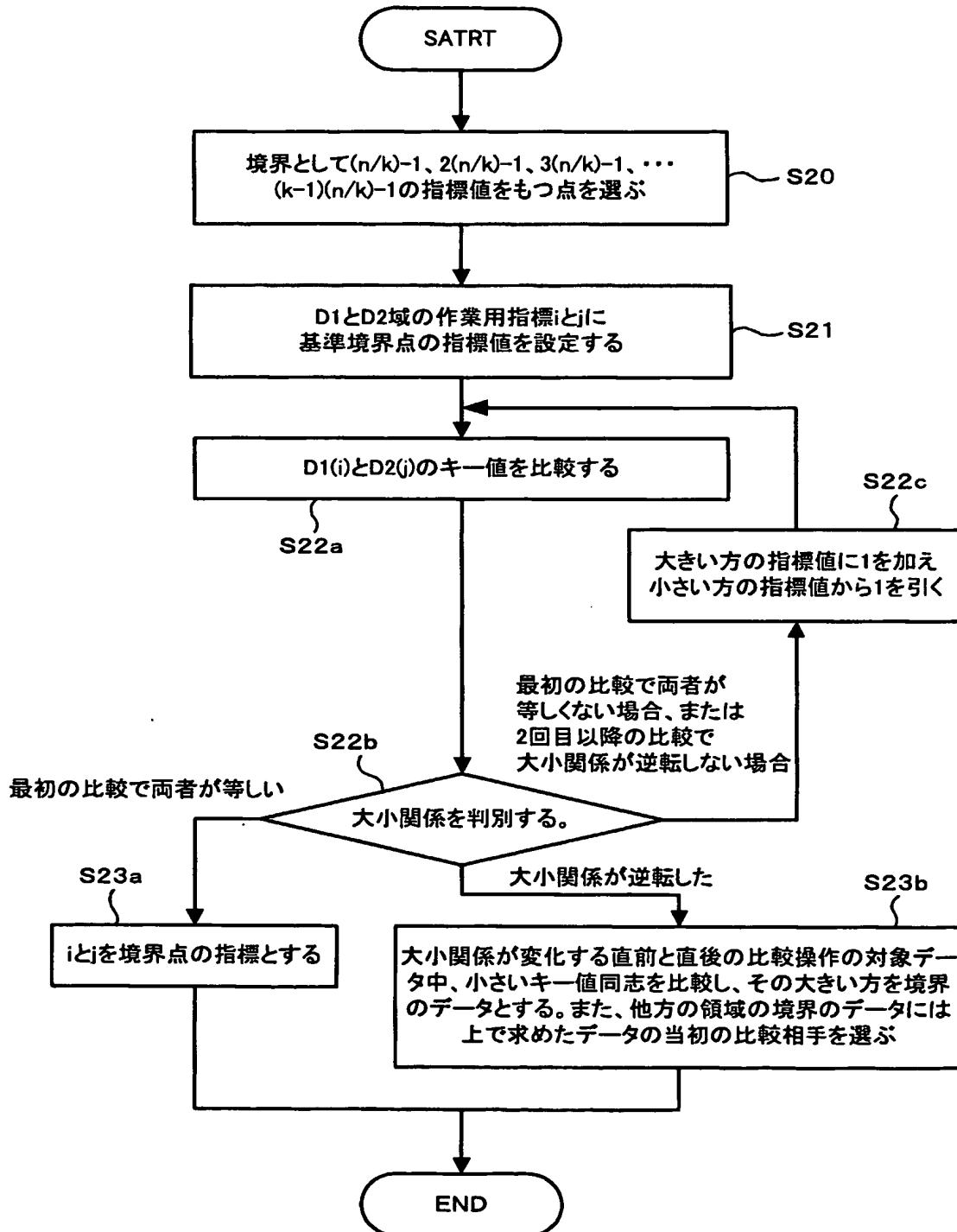
8 / 14

図 8



9 / 14

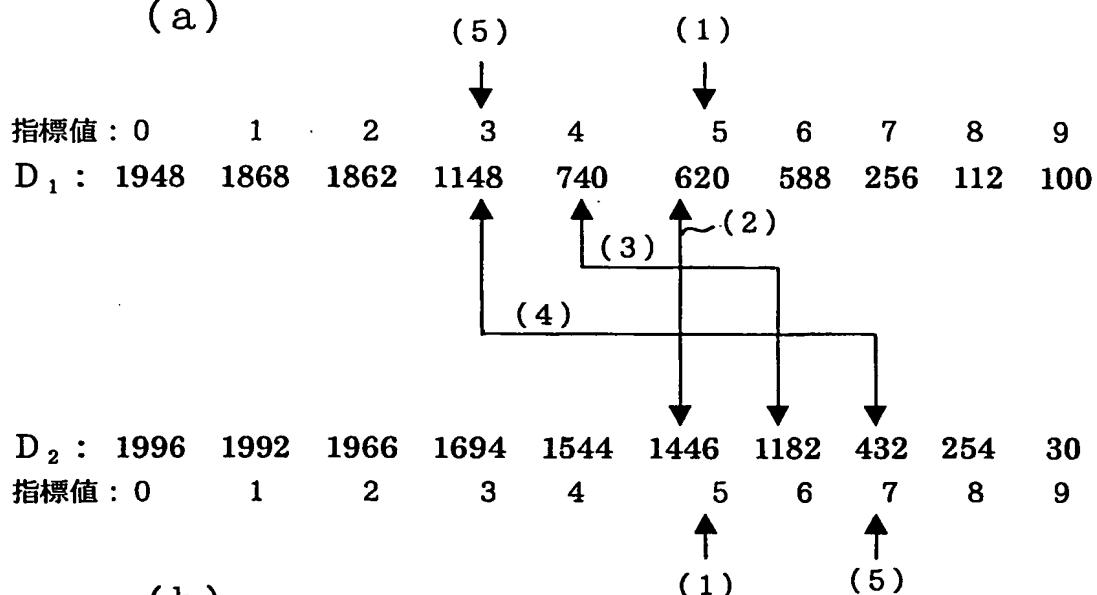
図 9



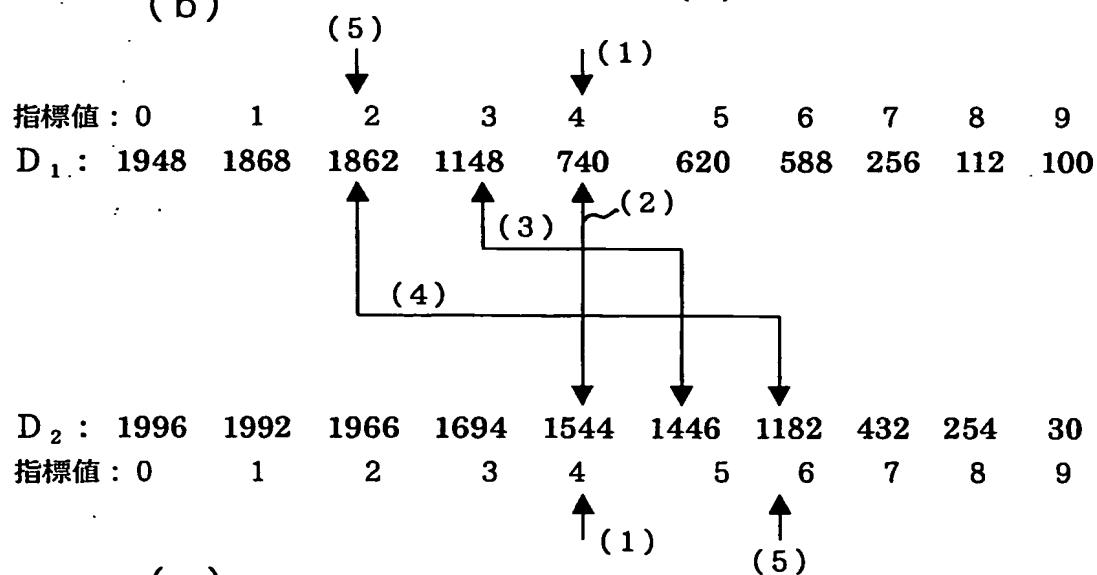
10 / 14

図 10

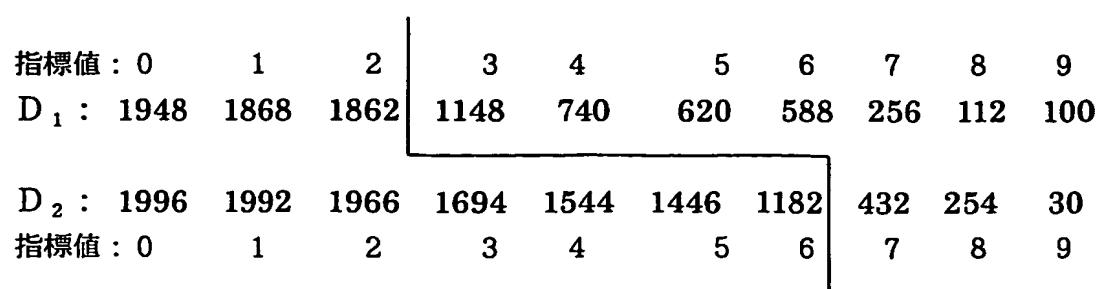
(a)



(b)



(c)



11 / 14

四 1 1

(a)

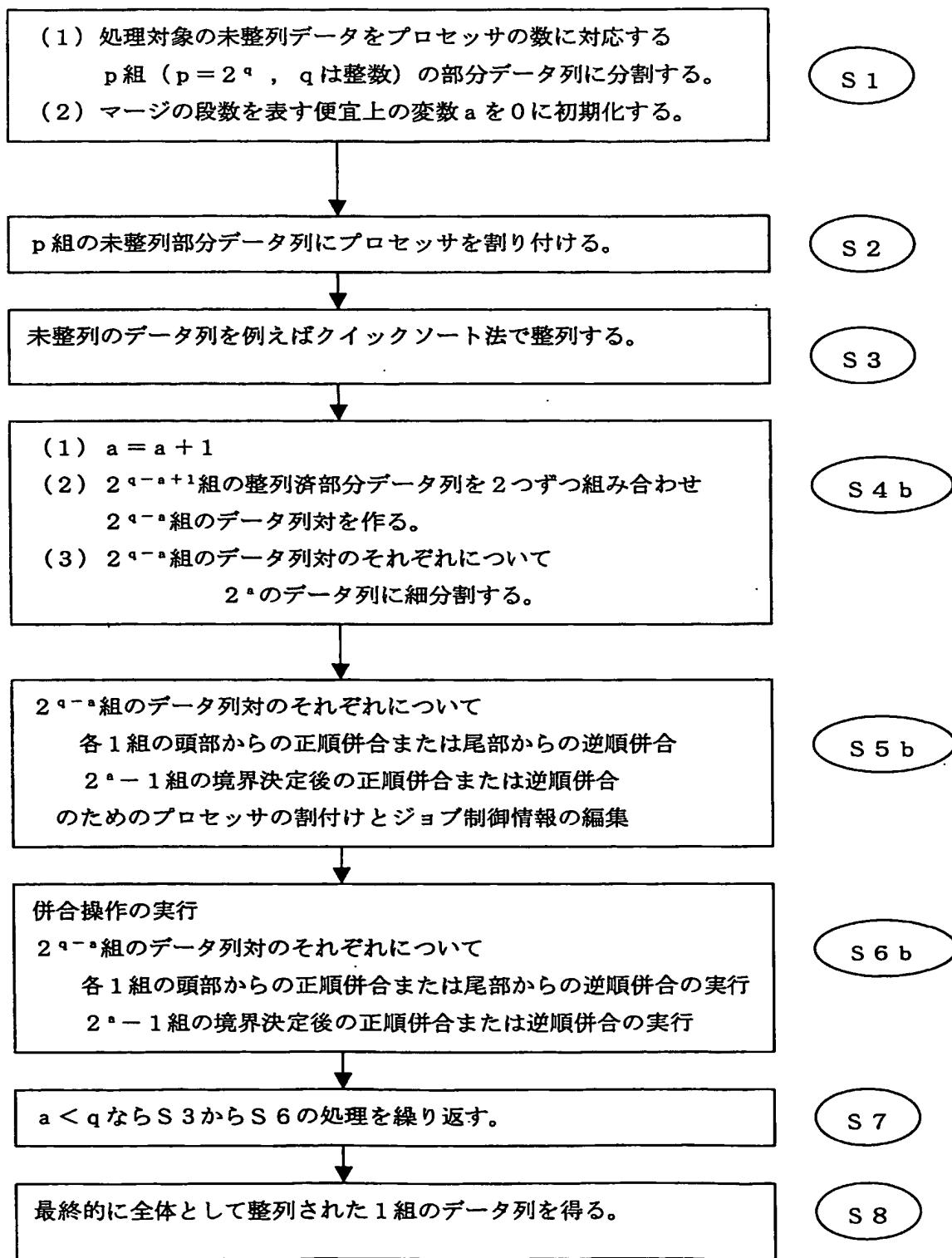
指標値 : 0	1	2	3	4	5	6	7	8	9	10	11	
D ₁ :	100	99	96	90	79	76	74	72	69	65	63	61
D ₂ :	98	97	89	86	84	82	81	80	77	73	67	66
指標値 : 0	1	2	3	4	5	6	7	8	9	10	11	

(b)

指標値 : 0	1	2	3	4	5	6	7	8	9
D ₁ : 100	98	97	93	90	89	88	87	86	60
D ₂ : 99	95	89	80	75	70	69	67	65	63
指標値 : 0	1	2	3	4	5	6	7	8	9

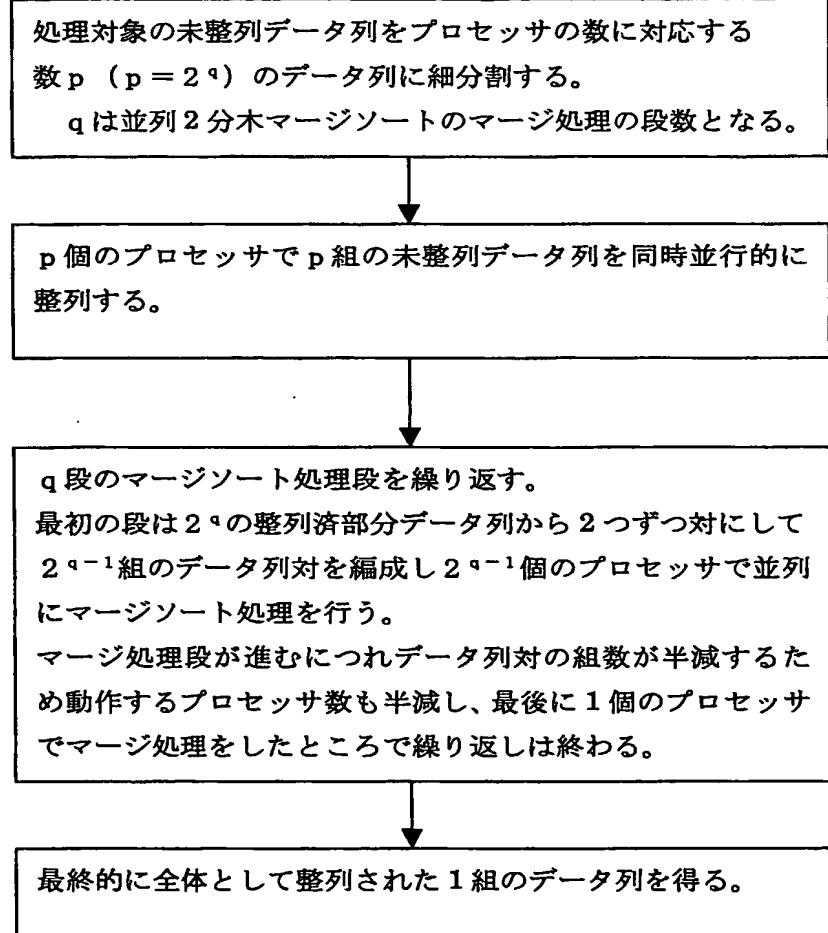
12 / 14

図 12



13 / 14

図 13



S100

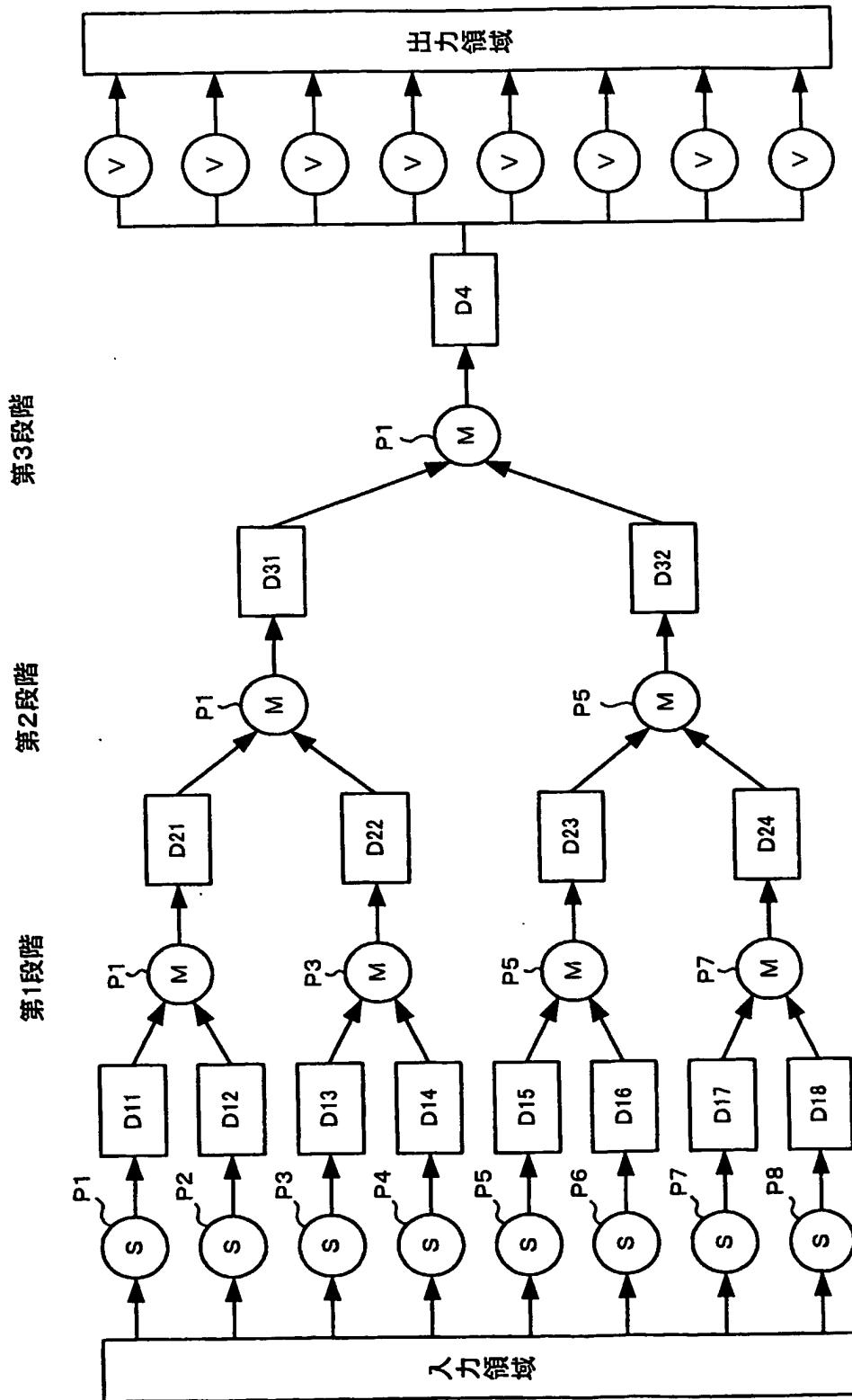
S101

S102

S103

14 / 14

図 14



INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP03/05319

A. CLASSIFICATION OF SUBJECT MATTER
Int.C1⁷ G06F7/24

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
Int.C1⁷ G06F7/24, G06F7/32

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
 Jitsuyo Shinan Koho 1922-1996 Toroku Jitsuyo Shinan Koho 1994-2003
 Kokai Jitsuyo Shinan Koho 1971-2003 Jitsuyo Shinan Toroku Koho 1996-2003

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2000-3342 A (NEC Corp.), 07 January, 2000 (07.01.00), Par. Nos. [0012], [0015] (Family: none)	1-11
A	US 5924093 A (Potter), 13 July, 1999 (13.07.99), Column 3, line 59 to column 4, line 28 (Family: none)	1-11
A	JP 1-288920 A (NEC Corp.), 21 November, 1989 (21.11.89), Claims (Family: none)	3

Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family
---	--

Date of the actual completion of the international search
25 July, 2003 (25.07.03)

Date of mailing of the international search report
12 August, 2003 (12.08.03)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Faxsimile No.

Telephone No.

A. 発明の属する分野の分類(国際特許分類(IPC))

Int. Cl' G06F 7/24

B. 調査を行った分野

調査を行った最小限資料(国際特許分類(IPC))

Int. Cl' G06F 7/24, G06F 7/32

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2003年
日本国登録実用新案公報	1994-2003年
日本国実用新案登録公報	1996-2003年

国際調査で使用した電子データベース(データベースの名称、調査に使用した用語)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	JP 2000-3342 A(日本電気株式会社), 2000.01.07(ファミリーなし), 段落0012, 段落0015	1-11
A	US 5924093 A(Potter), 1999.07.13(ファミリーなし), 第3欄第59行～第4欄第28行	1-11
A	JP 1-288920 A(日本電気株式会社), 1989.11.21(ファミリーなし), 特許請求の範囲	3

 C欄の続きにも文献が列挙されている。 パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの

「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献(理由を付す)

「O」口頭による開示、使用、展示等に言及する文献

「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」同一パテントファミリー文献

国際調査を完了した日

25. 07. 03

国際調査報告の発送日

12.08.03

国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)

郵便番号 100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官(権限のある職員)

田中 友章



5 E 9376

電話番号 03-3581-1101 内線 3520